

AIR PUBLICATION

1093F

VOLUME I

RADAR CIRCUIT PRINCIPLES

WITH AERIALS AND CENTIMETRE TECHNIQUE

*Prepared by direction of the
Minister of Supply*

A. C. Rowlands.

*Promulgated by order of the
Air Council*

J. H. Barnes.

AIR MINISTRY

First issued, June, 1946

Issued in Reprint, March, 1948

NOTE TO READERS

Air Ministry Orders and Vol. II, Part 1 leaflets either in this A.P., or in the A.P.'s listed below, or even in some others, may affect the subject matter of this publication. Where possible, Amendment Lists are issued to bring this volume into line, but it is not always practicable to do so, for example when a modification has not been embodied in all the stores in service.

When an Order or leaflet is found to contradict any portion of this publication, the Order or leaflet is to be taken as the over-riding authority.

When this volume is amended by the insertion of new leaves, the new or amended technical information is indicated by a vertical line in the outer margin. This line is merely to denote a change and is not to be taken as a mark of emphasis.

Each such leaf is marked in the top left-hand corner of the right-hand page with the number of the A.L. with which it was issued.

LIST OF ASSOCIATED PUBLICATIONS

<i>A.P. No.</i>	<i>Title</i>
1093C	Introductory survey of radar—Part I
1093D	Introductory survey of radar—Part II
1093E	Interservices radar manual

RADAR CIRCUIT PRINCIPLES
with aerials and centimetre technique

LIST OF CHAPTERS

CHAPTER 1—Circuits

CHAPTER 2—Radar aerials

CHAPTER 3—Transmission lines and waveguides

CHAPTER 4—Centimetre valves

CHAPTER 1

CIRCUITS

LIST OF CONTENTS

	Para.		Para.
Introduction	1	Comparison of the input impedance of the cathode follower and the RC coupled amplifier with resistive loads	
The superheterodyne receiver	6	Cathode follower	180
Bandwidth	8	Amplifier	184
Timing section and display	11	Resistance—capacitance oscillators	186
CR, LR and LCR circuits		Phase shift oscillator	187
Charge and discharge of a capacitor	18	Generators of square and pulse waveforms	
Discharge of the capacitor	21	The squarer	195
Growth of current in LR circuit	22	The multivibrator	206
Decay of current	24	The cathode coupled multivibrator	216
Frequency response of CR circuit	25	Delay circuits	219
Phasing	26	The flip-flop	220
Response of CR circuit to square and triangular waveforms	28	The single-valve delay circuit	223
Response of CR circuit to a square waveform	31	The blocking condenser	225
Asymmetrical square waveform	49	Production of pulses from square wave	227
Triangular waveform input	54	Production of a delayed pulse by means of a delay line	235
Short CR circuit	55	Cathode ray tubes	
Long CR circuit	56	Principles and functions	239
Differentiation and integration	57	Electrostatic focussing and deflection	242
CR circuit as a differentiating circuit	58	Electromagnetically-focussed and deflected tubes	262
Integrating circuits	61	Electrostatic focussing and electromagnetic deflection	271
CR circuit as an integrating circuit	63	Time base circuits for electrostatic deflection	
Tuned circuits	65	The single triode time base	276
Response of tuned circuit to an RF pulse	70	Use of anode-catching diode	285
Pulse shape from bandwidth considerations	71	Variation of run-down time and amplitude	289
Pulse shape from consideration of build-up and decay of oscillations	73	The output impedance of the Miller valve	293
Off-tune effects	79	Balanced output stages	296
Ringing circuits	80	Floating paraphase	300
Diode detection	87	Long-tailed pair	308
Sine wave input	90	Balanced shift control	313
RF pulse input	94	Linear time base circuits for electromagnetically-developed tubes	314
DC restoration and limitation		Pedestal waveform generator	318
DC restoration	99	Self-capacitance of deflector coils	319
Limitation	114	Push-pull output stage, with step-down transformer	320
Valve theory	121	Application of shift	322
Valve constants	124	Range amplitude display	
The equivalent circuit of a voltage amplifier	133	DC restoration	326
Load lines	138	Intensity modulation of CRT	
Anode bottoming	144	Plan position indicator (PPI) displays	330
Dynamic characteristics	145	Electrostatically-deflected PPI	331
Pentode valve as a constant current generator	148	Electromagnetically-deflected PPI	348
Resistance—capacitance (RC) coupled amplifier	149	Calibration	351
Frequency response of the RC coupled amplifier	152	Simple calibrator	354
Thevenin's theorem	154	Crystal controlled calibrators	357
Low frequencies	155	Strobe circuits	358
High frequencies	156	The super-regenerative receiver	361
Response to square waveform	159	Pulse modulation of transmitters	373
Edges of output waveform	160	Pulse transformers	381
Flat portions of the output waveform	162	Use of pulse transformers in conjunction with hard valve modulators	382
Biasing arrangements	164	Overswing diode	385
Inductance compensation	166	Driver stage	386
The cathode follower	169	Use of spark gaps for modulation	389
Amplification	170	Trigger circuit for trigatron	391
Handling capacity	171	Use of delay network to open spark gap	392
Biasing arrangements	173	Resonant choke charging	399
Output impedance	174	Constant current charging	401

LIST OF ILLUSTRATIONS

	<i>Fig.</i>		
Block diagram of a typical radar equipment	1	Positive limiting to earth potential by a parallel diode limiter	...
DC pulses	2	Negative limiting to earth potential	...
RF pulses	3	Positive limiting by a parallel diode circuit	...
Block diagram of superhet receiver	4	Negative limiting by a parallel diode circuit	...
Pulse waveforms from superhet receiver	5	Amplifier and parallel diode limiter	...
Waveforms	6	Limiter used for squaring	...
Bandwidths	7	Removal of positive or negative peaks from differentiated square waveform	...
Simple time base circuit and waveforms	8	Double-diode limiting	...
Typical arrangement of timing circuit with waveforms (1)	9	Characteristic curves of a triode (MH4)	...
Typical arrangement of timing circuit with waveforms (2)	10	Characteristic curves of a pentode (EF50)	...
Charge of a capacitor in an RC circuit (1)	11	Suppressor grid characteristics for pentode	...
Discharge of a capacitor in an RC circuit (2)	12	Equivalent circuit of a voltage amplifier	...
Growth of current in an LR circuit	13	Current-voltage graph for single resistor	...
Decay of current in an LR circuit	14	Current-voltage graph for two series resistors	...
Frequency response of CR circuit	15	Load lines	...
CR circuit, poor low response (2)	16	Anode bottoming and dynamic characteristic for anode load	...
CR circuit, poor high response	17	Pentode characteristic showing the region where VA has little control	...
Application of symmetrical square waveform to a short CR circuit	18	RC coupled amplifier	...
Application of symmetrical square waveform to medium CR circuit	19	Equivalent diagrams of RC amplifier for middle band of frequencies	...
Variation of V_0 in equilibrium state ($CR \approx T$)	20	Equivalent diagrams of RC amplifier for low frequencies	...
Application of symmetrical square waveform to "long" CR circuit	21	Equivalent diagrams of RC amplifier for high frequencies	...
Application of asymmetrical square waveform to CR circuit (1)	22	Frequency response curve for RC amplifier	...
Application of asymmetrical square waveform to CR circuit (2)	23	Distortion of pulse waveform by RC coupled amplifier	...
Application of asymmetrical square waveform to CR circuit (3)	24	RC amplifier with self-bias arrangement	...
Application of symmetrical triangular waveform to a "short" CR circuit	25	RC amplifier with unbypassed bias resistor	...
Application of symmetrical triangular waveform to a "long" CR circuit	26	Inductance compensation	...
Differentiated waveform	27	Basic cathode follower circuit	...
Integrating circuits	28	Equivalent circuit of cathode follower (1)	...
Parallel tuned circuits	29	Alternative equivalent circuit for cathode follower	...
Effect of resistance on the impedance frequency curve of a parallel-tuned circuit	30	Biassing arrangements for the cathode follower	...
Response curve of a parallel-tuned circuit	31	Equivalent circuit of cathode follower (2)	...
R.F. pulse shape across tuned circuit	32	Equivalent circuit of cathode follower with capacitive load	...
Derivation of a pulse shape	33	Distortion of positive pulses by shunt capacitance across R_K	...
Off-tune effects	34	Distortion of negative pulses by shunt capacitance across R_K	...
Ring circuit...	35	Cathode follower circuit	...
Ring of a resistanceless tuned circuit	36	Amplifier circuit	...
Ring of a damped tuned circuit	37	Basic circuit of the phase shift oscillator (PSO)	...
Ring of tuned circuit with pentode switching valve	38	RC phase-shifting network	...
Ring of tuned circuit by triode switching valve	39	3-mesh RC phase-shifting network	...
Production of narrow pulses from a damped ringing circuit	40	Phase shift oscillators with phase-advancing networks	...
Diode detector circuits	41	Phase shift oscillators with phase-retarding networks	...
Diode detector waveforms	42	Phase shift oscillator circuit	...
Diode detection of RF pulses	43	Circuit illustrating action of grid stopper	...
Filter outputs for diode detectors	44	Voltage waveforms of fig. 95	...
DC restorer waveforms (1)	45	Simple squarer stage without bias	...
DC restorer waveforms (2)	46	Waveforms of circuit of fig. 97	...
Square waveform applied to long CR circuit with R returned to voltage V	47	Modifications of waveforms of fig. 98 when no bottoming action occurs	...
Square waveform applied to negative DC restorer circuit	48A	Waveforms of squarer with negative bias	...
Square waveform applied to positive DC restorer circuit	48B	Waveforms of squarer with positive bias	...
DC restoration of asymmetrical square waveform	49	Circuit of squarer stage with arrangement for positive or negative bias	...
DC restoration of a sine waveform	50	Subsidiary circuit illustrating charging of coupling capacitor C	...
Negative DC restoration to potential V	51A	Effective circuit of fig. 103 when V_1 is suddenly cut off	...
Positive DC restoration to potential V	51B	Waveforms of circuits of figs. 103 and 104	...
Grid clamping circuit	52		

	Fig.		Fig.
Skeleton circuit of pentode multivibrator	106	Complete circuit diagram of Miller time base	160
Waveforms of circuit of fig. 106	107	Paraphase circuit	161
Variation of T_1 of fig. 107 by two methods	108	Floating paraphase circuit	162
Waveforms of circuit of fig. 106 when V_1 is not bottomed	109	Voltage dividers of floating paraphase	163
Circuit of cathode-coupled multivibrator	110	Voltage dividers with equal arms	164
Waveforms of circuit of fig. 110	111	Output impedance of floating paraphase	165
Illustrations of definition of "delay circuits"	112	Basic circuit of long-tailed pair	166
Circuit of flip-flop	113	Current changes in long-tailed pair	167
Waveforms of circuit of fig. 113	114	Typical long-tailed pair	168
Triggering of flip-flop from wide negative-going pulse	115	Equivalent circuit of stage-feeding coils	169
Triggering of flip-flop from wide positive-going pulse	116	Waveforms for fig. 169	170
Single-valve delay circuit	117	Typical tetrode output stage	171
Waveforms of circuit of fig. 117	118	Cathode follower output stage	172
Circuit of typical blocking oscillator	119	Equivalent circuit of cathode follower output stage	173
Waveforms of circuit of fig. 119	120	Cathode follower waveforms	174
Action of short CR circuit	121	Pedestal waveform generator	175
Basic circuit of "pip" eliminator stages	122	Waveforms for pedestal waveform generator	176
Waveforms of circuit of fig. 122 with negative bias	123	Push-pull output stage with step-down transformer	177
Waveforms of circuit of fig. 122 with positive bias	124	Application of shift current	178
Circuit for production of pulses by ringing circuit	125	Range amplitude display	179
Waveforms of circuit of fig. 125	126	Bright-up or black-out waveforms	180
Delay line showing four sections	127	Effect of DC restoration	181
Five sections of delay line showing construction	128	Input circuit to signal deflector plates	182
Circuit for production of pulses by delay line	129	Input circuit to deflector plates for balanced signals	183
Waveforms of circuit of fig. 129, when $R = Z_0$	130	Intensity modulation	184
Waveforms of circuit of fig. 129, when $R < Z_0$	131	Analysis of radial scan	185
Waveforms of circuit of fig. 129, when $R > Z_0$	132	Time base waveforms for electrostatic PPI	186
Circuit for production of delayed pulse	133	Magslip	187
Waveforms of circuit of fig. 133 with $R = Z_0$	134	Output from stator	188
Output waveform of circuit of fig. 131 when $R \neq Z_0$	135	Diode clamping circuit	189
Diagram of the elements of a CRT	136	Waveforms for diode clamping circuit	190
Gun and deflector plates of an electrostatically-focused and deflected CRT	137	Triode clamping circuit	191
Electrostatic focusing	138	Balanced time base waveform	192
Deflector plates	139	Mounting of single pair of deflector coils	193
Powe. supplies for an electrostatic CRT	140	Time bases produced for different positions of deflector coils	194
Astigmatism	141	Calibrator displays	195
Deflection defocus	142	Calibrator circuit	196
Trapezium distortion	143	Calibrator waveforms	197
Electromagnetic CRT	144	Strobe waveforms	198
Electromagnetic focusing	145	Strobe circuit and waveforms	199
Electromagnetic deflecting system	146	Build-up and decay of oscillations in super-regenerative detector	200
Deflector coils	147	Detection of output from super-regenerative stage	201
Fundamental time base circuit	148	Quench mush	202
Output waveform of ideal self-running time-base	149	Block diagram of super-regenerative receiver	203
Waveforms of ideal triggered time base	150	RF stage and detector	204
Circuit of single triode time base	151	Series switch modulation	205
Graphical method of finding V_{min}	152	Use of pulse transformer	206
Waveforms of circuit of fig. 151	153	Distortion produced by pulse transformer	207
Basic circuit of suppressor-triggered time Miller base	154	Overswing diode	208
Mutual characteristic of VR.91 for screen volts = 100	155	Possible driver stage	209
Waveforms of the suppressor-triggered Miller time base	156	Driver stage employing pulse transformer	210
Circuit conditions of Miller time base during run down	157	Trigatron	211
Circuit and waveforms of suppressor-triggered Miller time base	158	Trigger circuit for trigatron	212
Two methods of controlling amplitude duration of run-down	159	Waveforms of trigger circuit for trigatron	213
		Trigatron modulator circuit	214
		Charging circuit	215
		Equivalent circuit with trigatron closed	216
		Waveforms of trigatron modulator	217
		Resonant choke charging	218
		Resonant choke-charging circuit	219
		Resonant choke charging waveforms	220
		Constant current charging	221

CHAPTER 1

CIRCUITS

INTRODUCTION

1. The fundamental requirement of a radar equipment is to position an aircraft or other object accurately, i.e. to determine the range, azimuth and possibly elevation with respect to the radar equipment. A complete radar equipment, used to determine these three quantities, requires:—

- (i) A transmitter.
- (ii) A receiver.
- (iii) Aerial systems for transmitting and receiving.

A block diagram of such an equipment is shown in fig. 1.

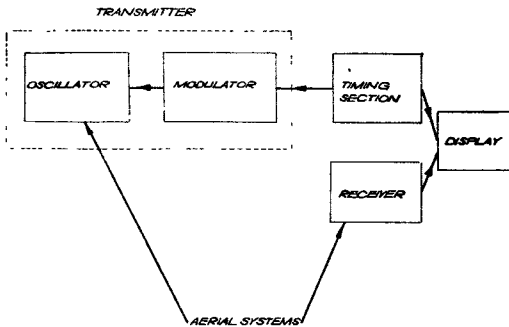


Fig. 1.—Block diagram of a typical radar equipment

2. The transmitter consists of:—

(i) *An oscillator.* Ideally, this should be a low power oscillator with good frequency stability, followed by RF amplifiers and an output stage. This is only possible at the lowest radar frequencies (CH, 20–60 Mc/s). Very much higher frequencies are now in general use, and amplification is not possible. It is usual to employ a single high-power oscillator.

20–60 Mc/s—Conventional valve oscillators used.

100–600 Mc/s—Tuned lines with conventional valves modified for VHF working.

3000 Mc/s and above—

Magnetrons and klystrons used.

(ii) *A modulator.*—The output waveform of this unit consists of DC pulses (fig. 2) which switch on the oscillator for the duration of the

pulses only, while the intervals between successive pulses the oscillator is non-conducting. The output of the oscillator thus consists of packets of RF oscillations or RF pulses, as shown in fig. 3.

Fig. 2.—DC pulses

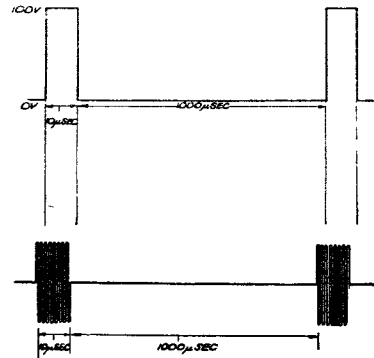


Fig. 3.—RF pulses

3. Receiver requirements are:—

- (i) High gain.
- (ii) Good signal/noise ratio.
- (iii) Faithful reproduction of pulse shape.

4. In most cases a superhet receiver is used to meet these requirements. It provides an efficient method of amplification, no other method being possible at frequencies of the order of 3000 Mc/s. Also, the pulses can be handled with little distortion if the receiver is well designed. Tuning is relatively simple, since the main tuning circuits are in the IF stages and are preset.

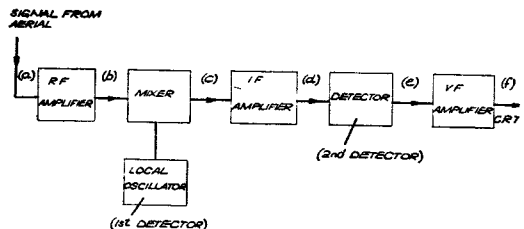


Fig. 4.—Block diagram of superhet receiver

5. Where extreme compactness is required, and faithful reproduction of pulse shape is not essential, a super-regenerative receiver is used,

e.g. in an IFF or beacon set. This receiver will be described in detail in a later section.

The superheterodyne receiver

6. A block diagram of a radar superhet receiver is given in fig. 4. Although adequate amplification of the signal can be obtained if it is fed directly to the mixer stage, the signal to noise ratio for carrier frequencies up to 600 Mc/s is improved by providing one or two stages of RF amplification. The signal to noise ratio depends mainly on the noise introduced by the first valve, and since at these frequencies amplifiers are less noisy than mixers, the RF stages are introduced. At higher frequencies, it is not possible to improve the signal/noise ratio by this means, so the mixer is made the first stage of the receiver.

7. In fig. 5, the output waveforms of the RF, IF, and VF stages are shown, when a $3 \mu\text{sec}$ RF pulse of carrier frequency 200 Mc/s is being received. A local oscillator signal of frequency 155 Mc/s is mixed with the amplified RF signal, to give an IF signal of frequency 45 Mc/s. This is amplified in the IF stages, converted into a DC pulse by the 2nd detector, then amplified by the VF stages before application to the deflector plates or grid of the display CRT. Detection is necessary since the mean voltage level of an RF pulse is zero.

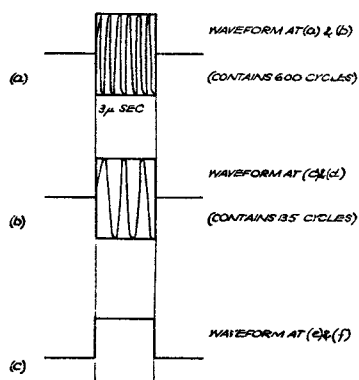


Fig. 5.—Pulse waveforms from superhet receiver

8. **Bandwidth.**—The receiver has to deal with pulse waveforms, which involves passing sharp edges, i.e. rapid changes in voltage, without distortion.

9. By Fourier analysis it can be shown that a periodic waveform consisting of ideal square DC pulses in which the changes of voltage level are instantaneous, as shown in fig. 6 (a), may be represented by an infinite series of sinusoidal components which are harmonics of the recurrence frequency. If such a waveform is to be passed without distortion, the receiver must respond equally to an infinite band of frequencies. This is impossible, but if some delay in the rise and fall of the pulse can be tolerated, then it is unnecessary for the infinitely high frequencies to be passed. Consider the case in which a delay of t secs. in the rise or fall of the pulse is permissible

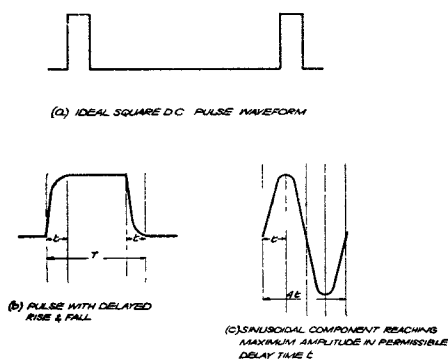


Fig. 6.—Waveforms

(fig. 6 (b)). By examination of figs. 6 (b) and 6 (c) it can be seen that for this to be realised, a sinusoidal component of frequency equal to

$\frac{1}{4t}$ c/s must be present. It is therefore neces-

sary for the stages of the receiver which handle pulses in their DC form, viz. the VF stages, to have a bandwidth extending from the prf to

at least $\frac{1}{4t}$ c/s if the pulse shape is to conform to the specified tolerance.

10. An RF pulse consists of the carrier frequency modulated by the band of sinusoidal components which constitute a DC pulse. It can be shown that this is equivalent to the carrier frequency together with sidebands of frequency equal to the carrier frequency plus or minus the modulating frequencies. Thus, an RF pulse of carrier frequency f_c c/s and having the same rate of rise and fall as the DC pulse considered above, consists of a range of frequencies extending from

$f_c - \frac{1}{4t}$ c/s to $f_c + \frac{1}{4t}$ c/s. The required bandwidth of the RF and IF stages is therefore $\frac{1}{2t}$ c/s, whilst that of the VF stages is $\frac{1}{4t}$ c/s (fig. 7).

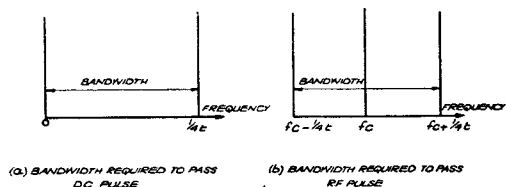


Fig. 7.—Bandwidths

Example

Time and rise of pulse $t = \frac{1}{10}$ μ sec.

VF bandwidth = $\frac{1}{4} \times 10^7$ c/s = 2.5 Mc/s.

IF and RF bandwidth = 5 Mc/s.

Timing section and display

11. A CRT is used to display the pulse and to enable the short delay times involved to be measured. A time-base is produced by moving the CRT spot across the screen at a uniform rate, starting at the instant the transmitter pulses. The time for complete traverse is made equal to the delay time of a pulse returning from a target at the maximum range which it is desired to display, e.g. to measure ranges up to 100 miles, the spot is made to move across the screen in 1 millisecond. The signal is displayed by deflection or intensity modulation of the time base trace. When an electrostatically deflected CRT is used, a linearly rising voltage waveform is required to produce the forward stroke or time base. The voltage must return to its initial level before the next pulse is transmitted. The manner of return is not very important, since it is not used for measurement purposes and is normally blacked-out.

12. The waveform is generally produced by charging and discharging a capacitor, the basic circuit being shown in fig. 8, together with waveforms. The capacitor C is allowed to charge so as to produce a linear rise of voltage for the required duration of the forward stroke; the discharge device is then caused to operate, discharging C to its original voltage, at which it remains until the start of the next time base. A switching or triggering voltage is required to make the discharge device

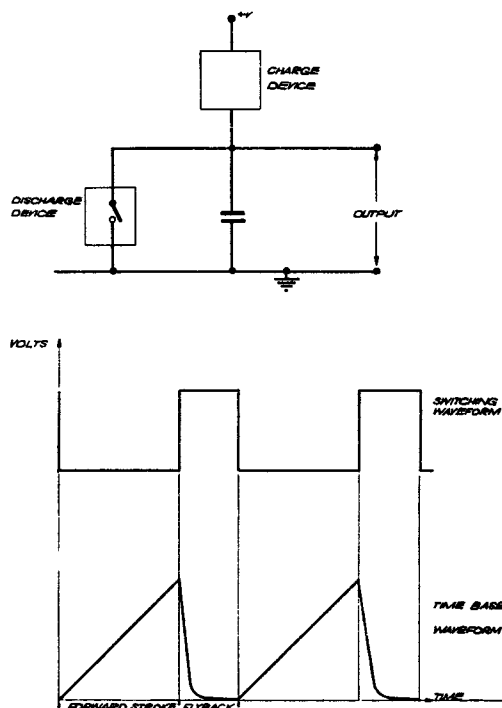


Fig. 8.—Simple time base circuit and waveform

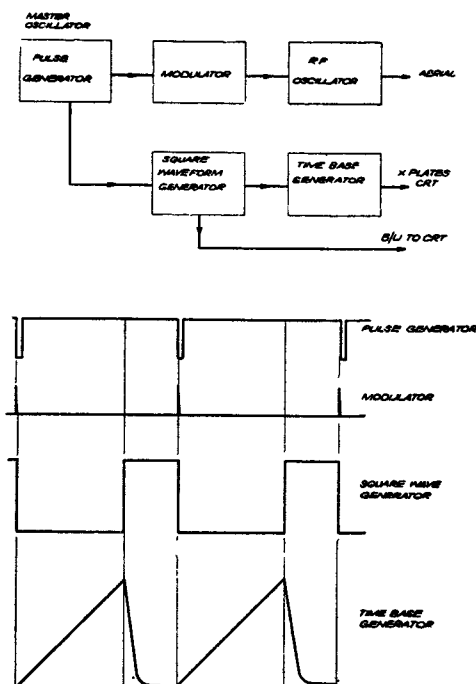


Fig. 9.—Typical arrangement of timing circuit with waveforms (1)

become alternatively conductive and non-conductive; this may take the form of a square waveform, positive going during the required conduction period, and negative for the non-conduction period, i.e. the forward stroke.

13. In order to brighten the trace during the forward stroke and extinguish the return, a square waveform of appropriate polarity during the forward stroke is applied to the grid or cathode of the CRT.

14. For different applications, time bases of various durations and with various ratios of forward stroke to flyback are required, and so it is necessary to provide circuits which will produce square waveforms with different relative widths of the positive and negative portions. The recurrence frequency of the square waveform and time base generators depends on the requirements of the equipment.

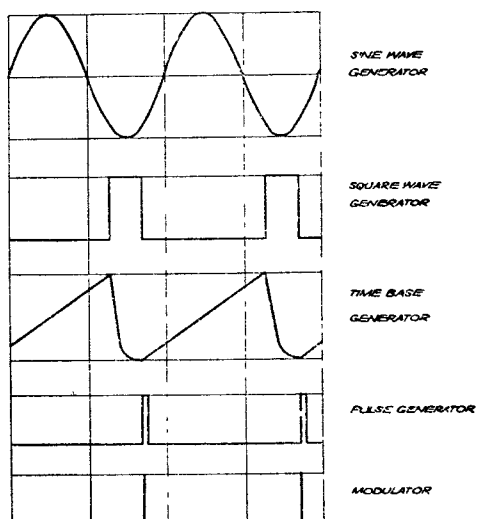
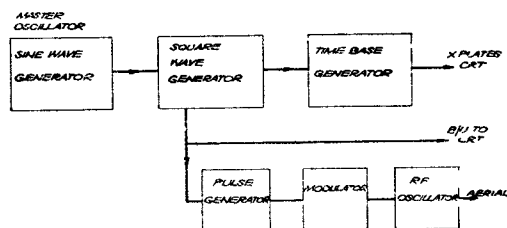


Fig. 10.—Typical arrangement of timing circuit with waveforms (2)

15. Another waveform to be generated in the timing section is the short DC pulse waveform required to trigger the modulator, so locking the transmission of the RF pulse to the start of the time base. Such a waveform has other important applications.

16. Figs. 9 and 10 show typical arrangements with waveforms of timing sections. In the block diagram of fig. 9, the master oscillator is a generator of short pulses. These are used to trigger the modulator, which produces pulses, locked to the triggering pulses, suitable for application to the RF oscillator. The pulses from the master oscillator also trigger a square waveform generator, the output of which is applied to the time base generator and determines the durations of the forward stroke and flyback of the time base waveform. The bright-up waveform for the CRT is taken from the square wave generator.

17. Fig. 10 shows an arrangement in which the master oscillator produces a sinusoidal waveform which is applied to a square waveform generator. The output of this stage triggers the time base generator, and is also applied to the pulse generator which feeds the modulator. Again, the bright-up waveform is taken from the square wave generator.

CR, LR AND LCR CIRCUITS

Charge and discharge of a capacitor

18. If a steady voltage V_0 is applied to a series CR circuit, see fig. 11, then, at any instant,

$$V_0 = V_c + R i$$

where V_c = voltage across C

and i = current flowing

the solution of which is

$$V_c = V_0 (1 - e^{-t/CR}) \dots \dots (1)$$

$$V_R = V_0 - V_c$$

where V_R = voltage across resistor.

At time $t = 0$

$V_c = 0$, i.e. the whole voltage V_0 appears across the resistor.

When $t = CR$

$$V_c = V_0 \left(1 - \frac{1}{e} \right)$$

$$= V_0 \left(\frac{e - 1}{e} \right) \quad (e \approx 2.7)$$

$$\approx 0.63 V_0$$

$$V_R \approx 0.37 V_0$$

The quantity CR is called the *time constant* of the circuit.

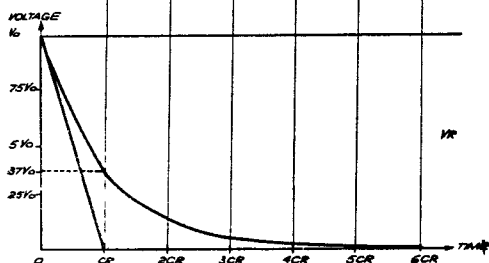
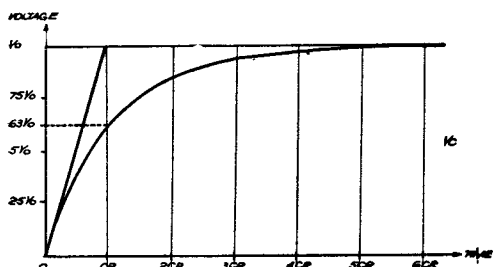
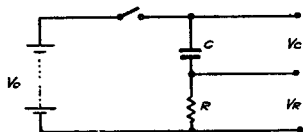


Fig. 11.—Charge of a capacitor in an RC circuit (1)

19. The percentage charge for various multiples of the time constant CR is given in the table below:—

$\frac{t}{CR}$	% charge of C
$\frac{1}{10}$	10%
1	63%
3	95%
4	98%
5	99.3%

Rate of change of voltage across C

$$= \frac{dV_c}{dt}$$

$$= \frac{V_0}{CR} e^{t/CR} \quad (\text{Differentiating equation (1)}).$$

$$\text{At } t = 0, \frac{dV_c}{dt} = \frac{V_0}{CR}$$

20. If the capacitor could continue to charge at this initial rate, a time CR would be required for V_c to rise to V_0 .

Thus the time constant CR can also be defined as the time taken by the capacitor to charge the applied voltage if the initial rate of rise of voltage is maintained (fig. 11).

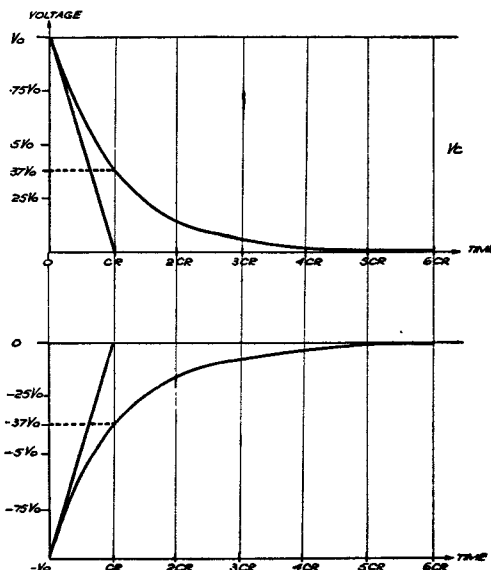
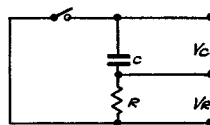


Fig. 12.—Discharge of a capacitor in an RC circuit (2)

Discharge of the capacitor

21. If, after the capacitor has completely charged to V_0 , the voltage source is shorted out, we have

$$Ri + V_c = 0 \text{ at any instant,}$$

the solution of which is,

$$V_c = V_0 e^{-t/CR}$$

At time $t = 0$, $V_c = V_o$

$V_R = -V_o$, since $V_R + V_c$
must be zero.

At time $t = CR$, $V_c = \frac{V_o}{e} \approx 0.37V_o$

Therefore $V_R \approx -0.37V_o$ since $V_R + V_c = 0$.

After time CR , voltage across the condenser has fallen to approximately a third of its initial value, see fig. 12.

Growth of current in LR circuit

22. The growth and decay of current in an inductor connected in series with a resistor, when a steady voltage is applied or removed, is a similar problem to that of the charge and discharge of a capacitor described above.

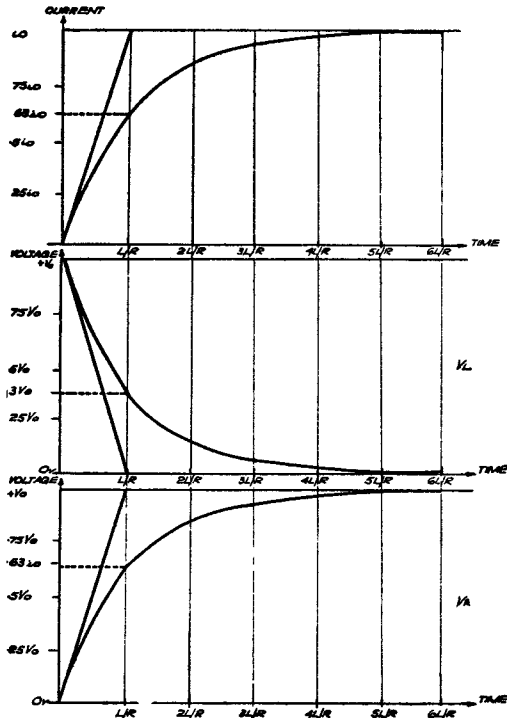
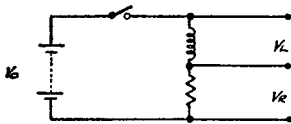


Fig. 13.—Growth of current in an LR circuit

23. If a steady voltage V_o is applied to the circuit (fig. 13), then at any instant

$$L \frac{di}{dt} + Ri = V_o \quad \text{where } i = \text{current flowing}$$

the solution of which is

$$i = i_o (1 - e^{-Rt/L})$$

where $i_o = \frac{V_o}{R} = \text{final steady current.}$

$$V_L = L \frac{di}{dt} = L \cdot \frac{R}{L} \cdot i_o \cdot e^{-Rt/L}$$

where $V_L = \text{voltage across the inductor.}$

$$= V_o e^{-Rt/L}$$

$$V_R = Ri = V_o - V_L$$

where $V_R = \text{voltage across resistor.}$

At time $t = 0$

$$i = 0; V_L = V_o; V_R = 0.$$

When $t = \frac{L}{R}$, the time constant

$$i = i_o (1 - e^{-1}) \approx 0.63 i_o \quad (e \approx 2.7)$$

$$V_L = V_o e^{-1} \approx 0.37 V_o$$

$$V_R \approx 0.63 V_o$$

The rate of rise of current $\frac{di}{dt}$ can be obtained from

$$L \frac{di}{dt} = V_o e^{-Rt/L}$$

$$\frac{di}{dt} = \frac{V_o}{L} e^{-Rt/L}$$

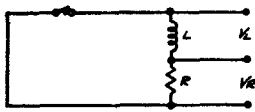
$$\text{When } t = 0, \frac{di}{dt} = \frac{V_o}{L}$$

$$= \frac{V_o}{R} \cdot \frac{R}{L}$$

$$= i_o \cdot \frac{R}{L}$$

Therefore the time constant $\frac{L}{R} = \text{time required}$

for the current to rise to the maximum value if the initial rate of growth is maintained.



When $t = 0$

$$i = i_0$$

$$V_L = -V_0$$

$$V_R = +V_0 \text{ since } V_L + V_R = 0.$$

When $t = L/R$

$$i = i_0 e^{-1}$$

$$\approx 0.37 i_0$$

$$V_L = -V_0 e^{-1}$$

$$\approx -0.37 V_0$$

$$V_R \approx 0.37 V_0$$

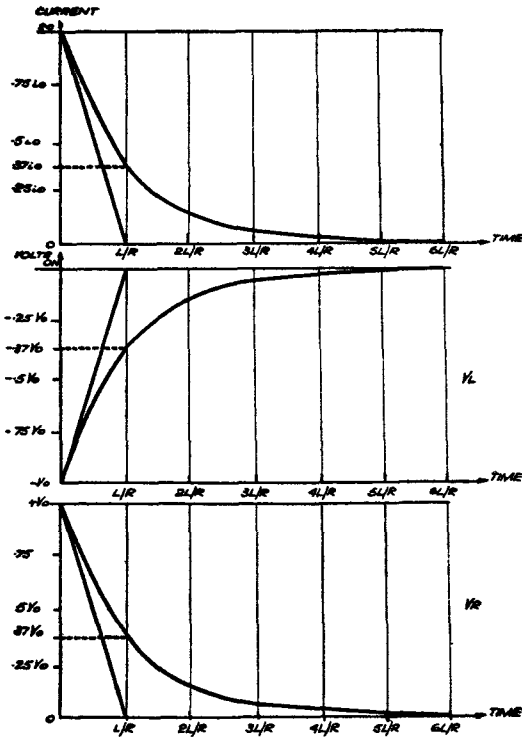


Fig. 14.—Decay of current in an LR circuit

Decay of current (fig. 14)

24. When the voltage V_0 is shorted out we have

$$L \frac{di}{dt} + Ri = 0$$

the solution being

$$i = i_0 e^{-Rt/L}$$

$$\frac{di}{dt} = -i_0 \frac{R}{L} e^{-Rt/L}$$

Hence

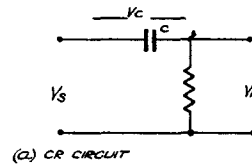
$$V_L = L \frac{di}{dt}$$

$$= -L \cdot \frac{V_0 R}{L} e^{-Rt/L}$$

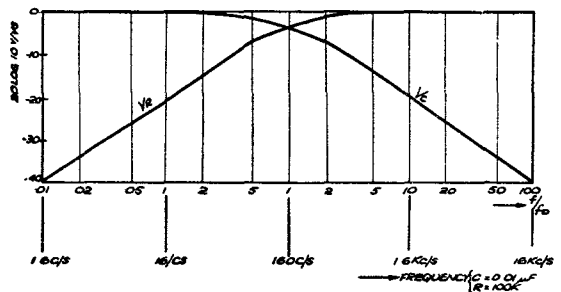
$$= -V_0 e^{-Rt/L}$$

Frequency response of CR circuit

25. A sine waveform is not distorted, but merely altered in amplitude and phase, when it is passed through a series CR circuit. The amplitudes of the voltages developed across C and R respectively in fig. 15 (a), are given by the expressions below, the variation in amplitude with the frequency of the input being



(a) CR CIRCUIT



(b) FREQUENCY RESPONSE CURVES

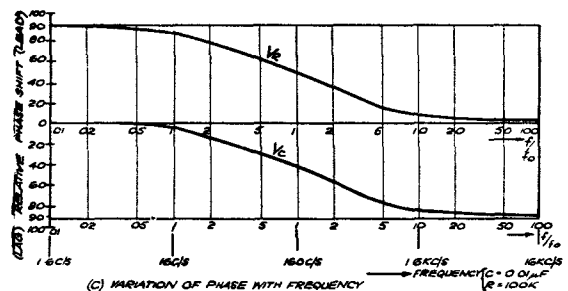


Fig. 15.—Frequency response of CR circuit

shown by the generalised frequency response curves of fig. 15 (b).

$$V_R = \frac{R}{\sqrt{R^2 + X_c^2}} \cdot V_s$$

where X_c = reactance of C.

$$= \frac{1}{\sqrt{1 + \left(\frac{X_c}{R}\right)^2}} \cdot V_s$$

$$V_c = \frac{X_c}{\sqrt{R^2 + X_c^2}} \cdot V_s$$

$$= \frac{1}{\sqrt{1 + \left(\frac{R}{X_c}\right)^2}} \cdot V_s$$

When the frequency is such that X_c is very much greater than R , then

$$V_c \approx V_s; V_R \approx 0.$$

As the frequency increases the reactance falls until at some frequency f_0 ,

$$X_c = R$$

$$\text{i.e. } \frac{1}{2\pi f_0 C} = R.$$

At this frequency f_0 , we have

$$V_R = \frac{1}{\sqrt{2}} \cdot V_s \\ \approx 0.707 V_s$$

$$V_c = \frac{1}{\sqrt{2}} \cdot V_s \\ \approx 0.707 V_s.$$

For further increase of frequency, the reactance of C falls off, ultimately becoming negligible in comparison with R . For these frequencies

$$V_R \approx V_s; V_c \approx 0.$$

Phasing

26. The variation in phase of V_R and V_c with respect to frequency is given by

$$\theta_R = \tan^{-1} \frac{X_c}{R} \quad \text{where } \theta_R = \text{phase angle of lead of } V_R \text{ with respect to } V_s.$$

$$\theta_c = \tan^{-1} \frac{R}{X_c} \quad \text{where } \theta_c = \text{phase angle of lag on } V_c \text{ with respect to } V_s.$$

At frequency f_0 where $X_c = R$

$$\theta_R = \theta_c \\ = \tan^{-1} 1 \\ = 45^\circ.$$

As the frequency is increased above f_0 , X_c falls, θ_R tends to zero and θ_c to 90° .

As the frequency is decreased below f_0 , X_c increases, θ_R tends to 90° and θ_c to zero.

27. The curves of fig. 15 (c) show this variation of phase with frequency. By comparison with the amplitude-frequency curves of fig. 15 (b), it can be seen that θ_R tends to zero over the same frequency range as V_R tends to V_s , and θ_c to zero as V_c tends to V_s .

Example

$$R = 100K$$

$$C = 0.01 \mu F$$

Frequency f_0 is given by

$$\frac{1}{2\pi f_0 \times 0.01 \times 10^{-6}} = 100 \times 10^3$$

$$f_0 = \frac{1}{2 \times 3.142 \times 10^{-8} \times 10^5} \text{ c/s} \\ \approx 160 \text{ c/s.}$$

The response of the circuit may now be obtained directly from the generalised curves of fig. 15 (b) and (c) by converting the frequency ratio f/f_0 scale to a scale of actual frequencies as shown in the diagram.

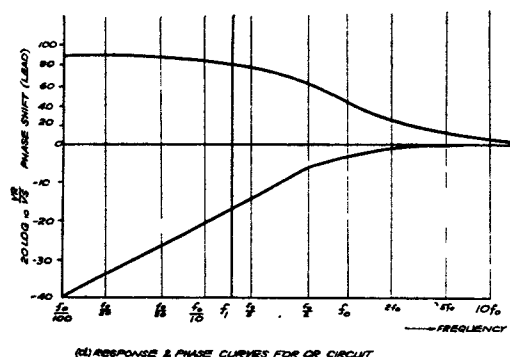
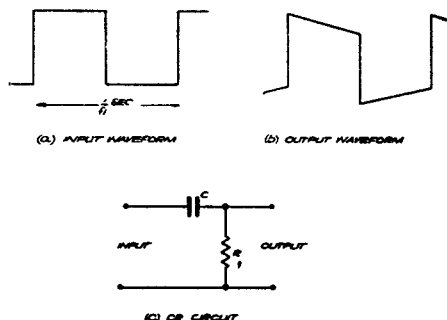
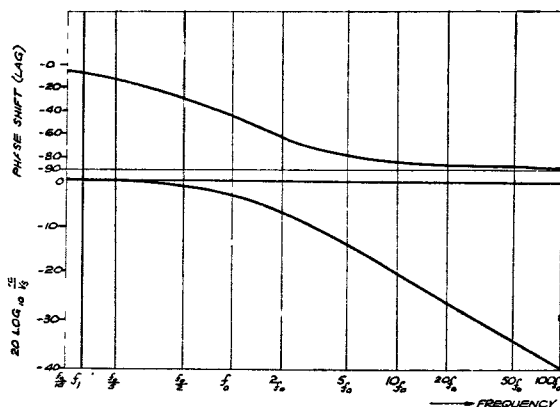
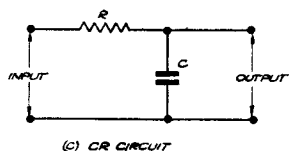
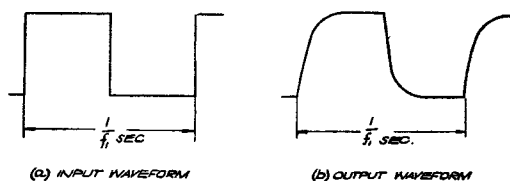


Fig. 16.—CR circuit poor low response

Response of CR circuit to square and triangular waveforms

28. Square pulse and triangular waveforms can all be Fourier analysed into series of sinusoidal harmonic components, the relative amplitudes of these components depending on the type of waveform. If a square waveform of recurrence frequency f_1 c/s is applied to a circuit which discriminates in amplitude and phase against its lower harmonic constituents, then it can be shown that the steady portions of the waveform are distorted in the output as shown in fig. 16 (a) and (b). Such an action occurs in a CR circuit in which the output is taken across the resistor and in which the CR values are such that it has an inadequate low frequency response, as indicated in fig. 16 (c) and (d).

29. If the circuit is such that it discriminates against the higher harmonics, analysis shows that in the output waveform the changes in voltage level are no longer instantaneous (fig. 17 (a) and (b)). This action takes place when the output is taken across the capacitor of a CR circuit which is deficient in high frequency response (fig. 17 (c) and (d)). A



(a) RESPONSE & PHASE CURVES FOR CR CIRCUIT

Fig. 17.—CR circuit poor high response

similar distortion occurs with a pulse or triangular waveform.

30. It is difficult, however, to determine accurately the shape of the output waveform on the basis of frequency response; an easier method is to consider the charge and discharge of the capacitor through the resistor produced by variation of the input voltage.

Response of CR circuit to a square waveform

31. We shall consider an ideal symmetrical square waveform of period $2T$ sec. being applied to a series CR circuit, and draw the voltage waveforms across each component for the following cases:—

- (i) $CR \leq T$
- (ii) $CR \approx T$
- (iii) $CR \geq T$

simplifying the treatment in each case by assuming that a resistanceless square wave generator is used.

$CR \leq T$. Assume C is initially uncharged.

32. When the input voltage is instantaneously raised from 0 to $+V$ volts, the total voltage change appears across R since capacitor C cannot charge instantaneously. C now begins to charge exponentially on the time constant CR and as the voltage V_c across the capacitor rises, V_R , the voltage across the resistor, falls, since at any instant the sum of the voltages across C and R must be equal to the applied voltage $+V$. C can charge almost completely in a time $5CR$, and since $CR \leq T$, then long before the input voltage is suddenly dropped to zero again, V_c attains the value V and V_R returns to zero.

33. When the input voltage is dropped to zero, the potential of the left-hand plate of C must fall with it, and since C cannot discharge instantaneously the potential of the right-hand plate is dropped by the same amount, i.e. from 0 to $-V$. This means that we have a voltage $+V$ across the capacitor and a voltage $-V$ across the resistor, the two adding up to the applied input voltage, i.e. 0 volts. C now discharges through R , V_c and V_R , returning to zero in a complementary manner since at any instant $V_c + V_R = 0$. The whole discharge takes a time approximately equal to $5CR$, and so is completed in a small fraction of this portion of the cycle. The whole action described is repeated during successive cycles of the input waveform.

34. It will be seen from the above discussion, that the voltage waveform across C is a distorted square wave, and that across R consists of a series of positive and negative "pips". If the

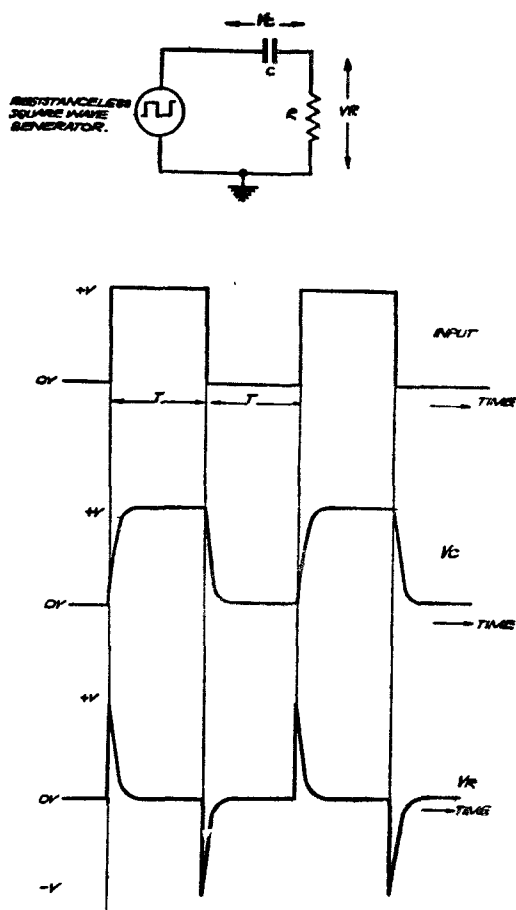


Fig. 18.—Application of symmetrical square wave to a short CR circuit

CR value is very small compared with T the capacitor waveform deviates little from the input waveform, and the pips produced across R are very narrow. As CR is increased the capacitor waveform becomes more distorted and the width of the pips increases.

35. "Short CR" circuits are used extensively in conjunction with pip eliminating valves (to be discussed later) to produce positive or negative pulses from square waveforms.

$CR \approx T$

36. Consider values of C and R which give a time constant such that the capacitor can charge or discharge through 70% of the applied voltage in the time T. Steady state conditions are not reached after one cycle of the applied waveform as for the short CR

circuit, and in order to simplify the problem of finding the voltage levels in the steady state, we will take an input square wave with the voltage levels 0 volts and 100 volts, see fig. 19.

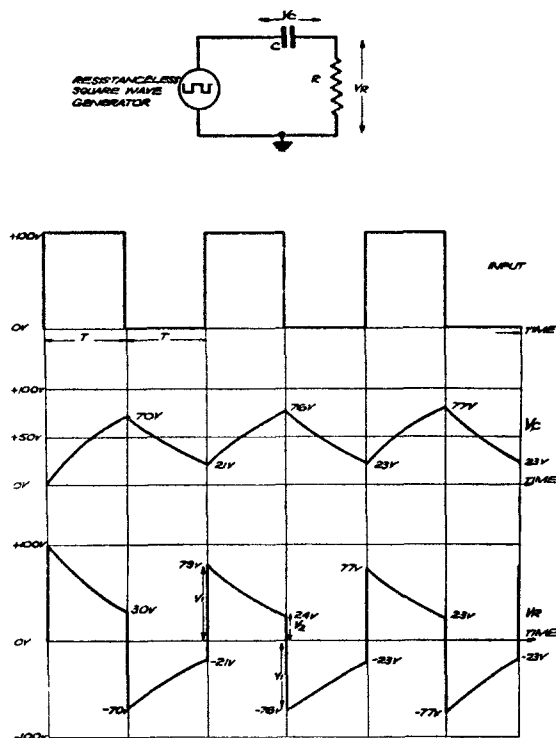


Fig. 19.—Application of symmetrical square waveform to medium CR circuit

37. Assume that initially C is uncharged. When the input voltage rises from 0 volts to 100 volts V_R rises instantaneously to 100 volts. Then C begins to charge and in time T, V_c rises to 70% of 100 volts, i.e. to 70 volts. Meanwhile V_R drops to 30 volts, at every instant the sum of the voltages across R and C being equal to 100 volts.

38. When the input voltage drops to 0 volts the left-hand plate of C falls in potential from 100 volts to 0 volts, and since C cannot discharge instantaneously, the voltage of the right-hand plate (i.e. V_R) drops from 30 volts to -70 volts, making the sum of V_c (+70v) and V_R (-70v) equal to zero. C then discharges during this portion of the input waveform. The initial voltage across C is 70 volts and in time T, C discharges through 70% of this voltage, i.e. V_c drops from 70 volts

to 70 - 49 volts = 21 volts. Meanwhile, V_R changes from -70 volts to -21 volts in a complementary manner.

39. In the next positive portion, the initial voltage across C is +21 volts and so C starts charging from 21 volts towards 100 volts, i.e. the initial effective applied voltage is $100 - 21 = 79$ volts. In time T, C charges through $70/100 \times 79$ volts ≈ 55 volts, i.e. V_c rises to $21 + 55 = 76$ volts.

40. Since $V_c + V_R = 100$ volts during this portion of the cycle, V_R jumps instantaneously from -21 volts to +79 volts and then falls exponentially from this value to $100 - 76 = 24$ volts.

41. Voltage levels for successive cycles of V_c and V_R are marked on the waveforms of fig. 19. The table below shows the changes of V_c which occur during the first three cycles:—

Cycle	Increase of V_c during positive input	Decrease of V_c during zero input
1	70 volts	49 volts
2	55 volts	53 volts
3	54 volts	54 volts

42. From this table it can be seen that the amounts of increase of V_c during successive positive half-cycles become progressively less; on the other hand the amounts of decrease of V_c during half-cycles in which the input is zero become larger. The steady state (i.e. that in which each cycle of the waveform across C or R is identical with the preceding one) is reached when the increase and decrease of V_c in successive half-cycles become equal. In the example, this stage is reached in the third cycle.

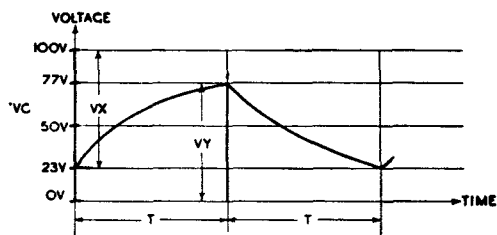


Fig. 20.—Variation of V_c in equilibrium state ($CR=T$)

43. Fig. 20 shows the variation of V_c in the equilibrium state. The effective applied voltage at the beginning of the charge period is equal to V_x and when the input voltage is removed, C discharges from a voltage V . Since the periods allowed for charge and discharge are equal, the rise and fall of V_c are the same percentages of V_x and V_y ,

respectively (i.e. 70% in the example considered). In order, then, to make the rise and fall equal in magnitude, it is necessary that $V_x = V_y$. Thus the maximum and minimum levels of V_c are symmetrically disposed with respect to the mean of the input voltage, i.e. $100/2 = 50$ volts. Since the mean level of V_c is 50 volts, that of V_R is 0 volts.

$$CR \gg T$$

44. Consider that values of C and R are chosen which enable the capacitor to charge or discharge by 10% in the half-period T of the input—again a symmetrical 0—100 volts square waveform.

45. The same building up of voltage across the capacitor as was described for the last case occurs here, the steady state again being reached when the amounts of charge and discharge of the capacitor during alternate half-periods of the input waveform are equal. However, it will be seen from the waveforms of fig. 21, that in this case a considerable number of cycles must occur before the steady state is reached.

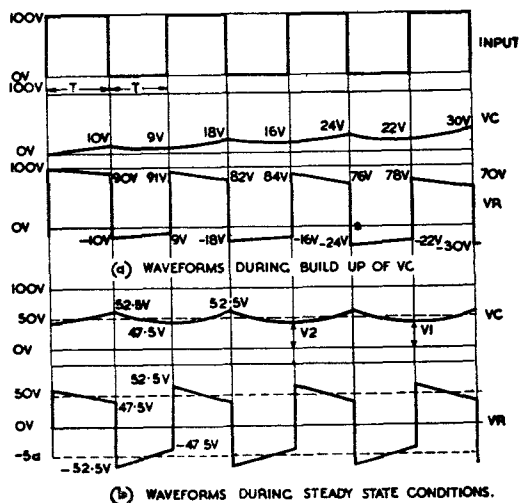
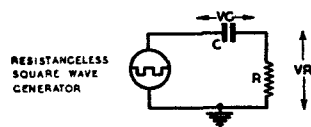


Fig. 21.—Application of symmetrical square waveform to "long" CR circuit

46. Fig. 21 (b) shows the variations of V_c and V_R during steady state conditions. The variation of V_c is symmetrical about the mean level of the input as in the last case, consequently V_1 and V_2 , the extreme values of

V_c are such that

$$V_1 + V_2 = 100 \text{ volts.}$$

Since 10% discharge occurs in time T

$$V_2 = .9V_1$$

$$1.9V_1 = 100 \text{ volts}$$

$$V_1 = 52.5 \text{ volts}$$

$$V_2 = 47.5 \text{ volts}$$

i.e. V_c varies by ± 2.5 volts about the mean voltage of the input, being approximately triangular in shape; V_R is a slightly distorted square waveform with the mean level at 0 volts, and the tops and bottoms varying by 5 volts during each half-cycle

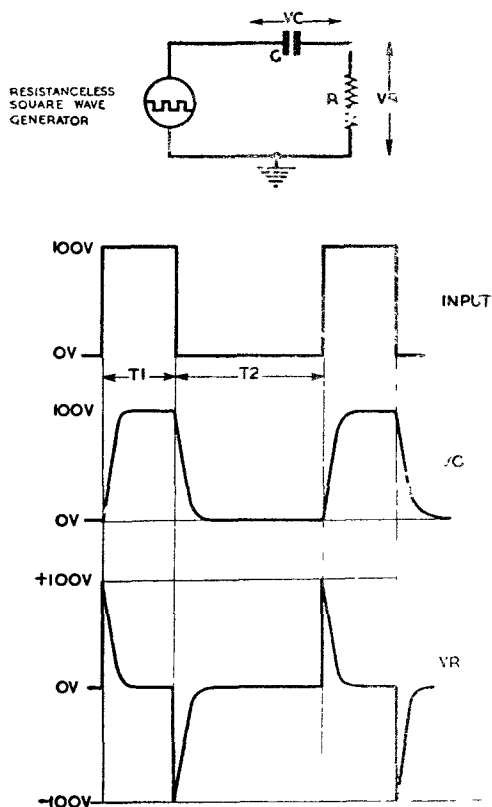


Fig. 22.—Application of asymmetrical square waveform to CR circuit (1)

47. The longer CR is made in comparison with T , the smaller becomes the variation in the voltage of capacitor C , which means that the waveshape across R becomes more nearly a replica of the input waveform. The time taken to reach the steady state increases with CR.

48. This is an important practical case, since it indicates that for square waveform to be passed without distortion by CR coupling circuits, the values of resistance and capacitance must be chosen to give a time constant which is very much longer than the recurrence period of the waveform to be passed.

Asymmetrical square waveform

49. Now consider the application of an asymmetrical square waveform, i.e. one with duration of positive portion (T_1) and negative portion (T_2) unequal. Waveforms are drawn for $T_1 < T_2$.

50. Fig. 22 shows the waveforms for $CR \ll T_1 < T_2$. Here the capacitor has

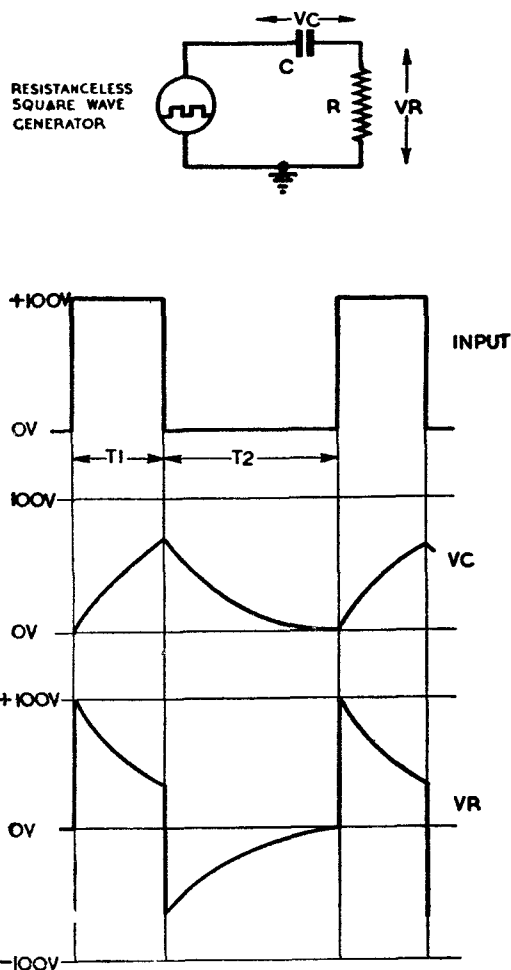


Fig. 23.—Application of asymmetrical square waveform to CR circuit (2)

time to charge and discharge completely, so a distorted version of the input waveform appears across C , while V_R consists of positive and negative pips.

51. Fig. 23 shows the waveform for a medium CR , e.g. $T_1 < CR$, $T_2 = 5CR$. C will only charge to a fraction of the applied voltage in time T_1 , but it discharges practically to zero in time T_2 .

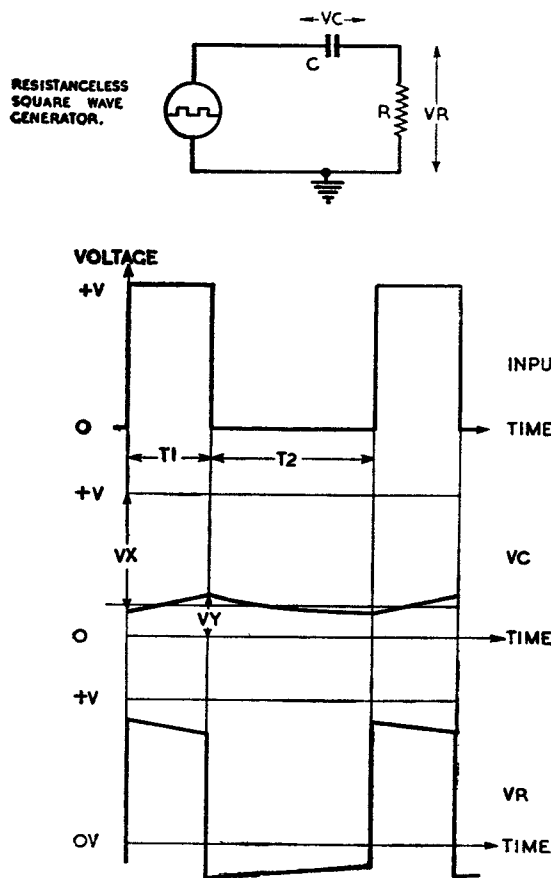


Fig. 24.—Application of asymmetrical square wave to CR circuit (3)

52. The waveforms of fig. 24 are drawn for $CR \geq T_2 > T_1$, so that little charge or discharge occurs. Steady state conditions again correspond to equal charge and discharge of C during the positive and negative portions of the cycle respectively. If the variation of V_c is negligible, the mean levels may be determined as follows.

Let V_x = mean effective applied voltage during positive portion

V_y = mean level of V_c .

The charge rate of C is proportional to V_x , the discharge rate to V_y .

$$\therefore V_x T_1 = V_y T_2$$

$$\text{i.e. } \frac{V_x}{V_y} = \frac{T_2}{T_1}$$

But the mean level of the input is such that the ratio of the excursions above and below it is also equal to T_2/T_1 . Hence C charges to the mean level of the input, and consequently the mean level of V_R is zero.

53. In the above treatment of the action of a CR circuit on a square waveform, it has been assumed for convenience that the input voltage variations are between zero and a positive voltage. If the input varies between positive and negative levels, the only difference produced is the alteration in the mean level of V_c with that of the input.

Triangular waveform input

54. The case of a symmetrical triangular waveform input to the circuit will be discussed. Consider that the input V_s takes the form shown in fig. 25, and assume that before the application of V_s , C is uncharged.

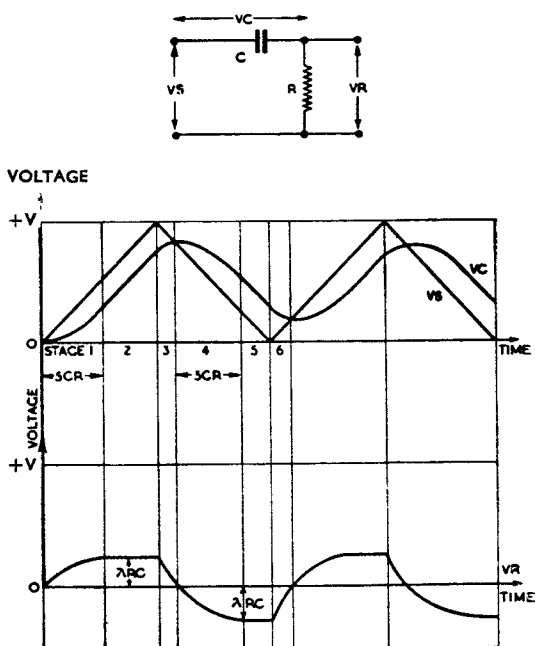


Fig. 25.—Application of symmetrical triangular waveform to a "short" CR circuit

55. Short CR circuit

Stage 1

At time $t = 0$, $V_s = 0$

$$V_c = 0$$

$$V_R = 0$$

Rate of rise of $V_c = 0$.

When V_s rises, C begins to charge, slowly at first since the effective applied voltage $V_s - V_c$ ($= V_R$) is small. As $V_s - V_c$ becomes larger, the rate of rise of V_c increases until eventually it becomes equal to the rate of rise of V_s . It can be shown that this occurs in a time approximately equal to $5CR$.

Stage 2

After a time $5CR$, V_c rises at the same rate as V_s and $V_R (= V_s - V_c) = \text{constant}$.

Stage 3

When the phase of the input reverses, C continues to charge so long as V_s exceeds V_c , but as the effective applied voltage ($= V_R$) decreases to zero, so does the charge rate fall to zero.

Stage 4

Beyond this point V_s falls below V_c , the effective applied voltage becomes negative and C begins to discharge. The discharge rate increases until it equals the rate of fall of V_s , in a similar manner to that which occurs for the positive going half cycle.

Stage 5

V_c falls at the same rate as V_s , and $V_R = \text{constant}$.

Stage 6

When the phase of the input again reverses, the action is similar to that described in stage 3.

If the CR value is made shorter, V_c becomes more nearly equal to V_s , and V_R becomes smaller in amplitude and approximates to a square waveform.

Long CR circuit

56. If the CR value of the circuit is so long that $5CR \gg T$, then Stage 1 is not completed before the input reverses polarity, as shown in fig. 26(a). Steady state conditions are not reached until a number of cycles have been completed, as in the case of a square waveform applied to a long CR circuit. Waveforms for the steady state are shown in fig. 26(b), from which it may be seen that the variation is

relatively small, and consequently V_R is a distorted reproduction of the input. The longer the CR , the closer do the shape and amplitude of V_R approximate to those of the input.

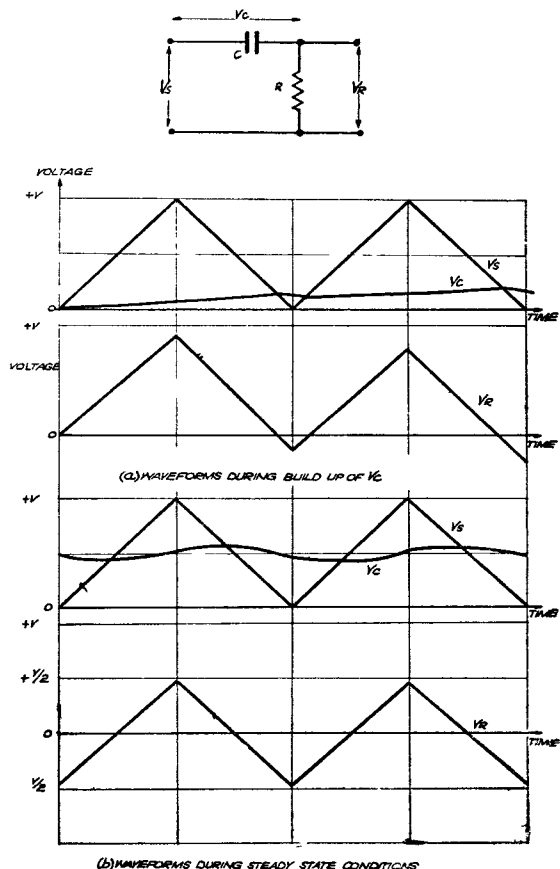


Fig. 26.—Application of symmetrical triangular waveform to a "long" CR circuit

Differentiation and integration

Differentiating circuits

57. A differentiating circuit is one in which the output is proportional to the differential coefficient or rate of change of the input, e.g. if the input and output voltages are V_s and V_o respectively, then

$$V_o = k \frac{dV_s}{dt} \text{ where } k \text{ is a constant.}$$

(i) Square waveform input

$$\frac{dV_s}{dt} = +\infty \text{ for the positive instantaneous voltage changes.}$$

$\frac{dV_s}{dt} = -\infty$ for the negative instantaneous voltage changes.

$\frac{dV_s}{dt} = 0$ for the constant voltage levels of the input.

Thus, a differentiated square waveform

$(V_o = k \frac{dV_s}{dt})$ consists of a series of positive

and negative spikes of zero width, and infinite amplitude (fig. 27a).

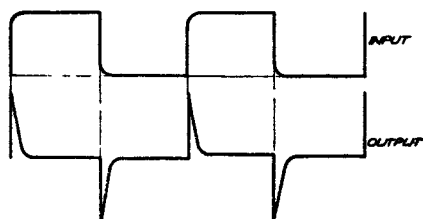
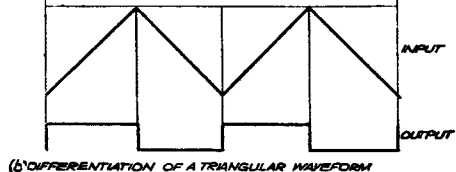
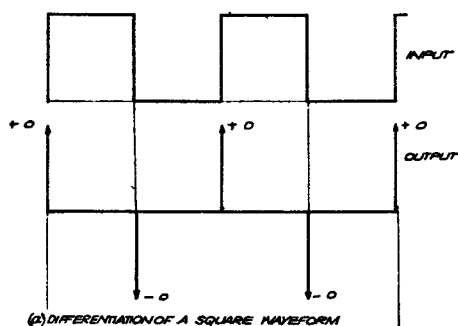


Fig. 27.—Differentiated waveform

(ii) Triangular waveform input

When the input voltage is rising at a uniform rate, $\frac{dV_s}{dt}$ is a constant positive quantity; similarly, when the input is falling steadily,

$\frac{dV_s}{dt}$ is a constant negative quantity. Hence the differentiated triangular voltage $(V_o = k \frac{dV_s}{dt})$ is a square waveform (fig. 27 (b)).

CR circuit as a differentiating circuit

58. It has been seen that when a square waveform is applied to a short CR circuit, the waveform across the resistor takes the form of narrow pips. As the CR value is made shorter, the pips become narrower and approximate more closely to the ideal differentiated square waveform described above, although the amplitude of each pip does not exceed the peak-to-peak amplitude of the input.

59. Again, when a triangular input is applied to a short CR circuit, the voltage variation across the resistor is a distorted square waveform which approximates to the differentiated triangular input.

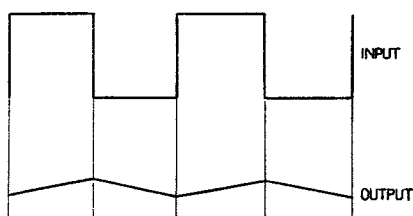
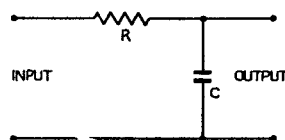
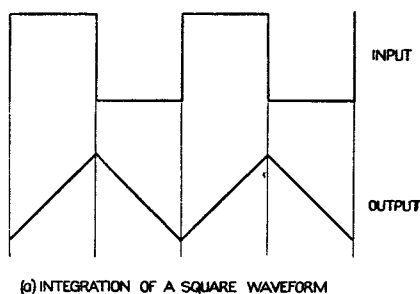
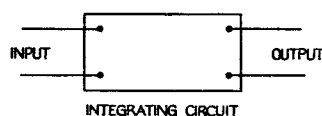


Fig. 28.—Integrating circuits

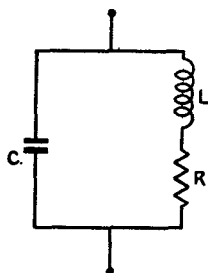
60. Consequently, a short CR circuit in which the output is taken across the resistor is commonly called a differentiating circuit.

Integrating circuits

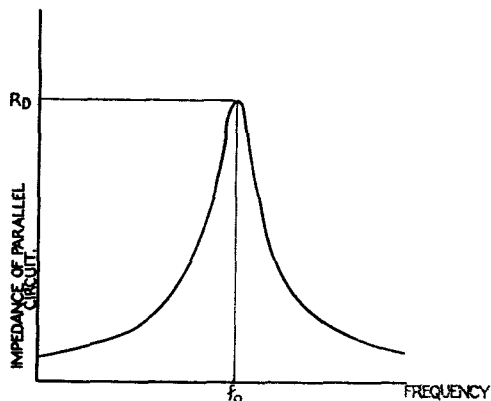
61. An integrating circuit is one in which the output is proportional to the integral of the input waveform. If the input and output voltages of an integrating circuit are V_s and V_o respectively, then

$$V_o = k \int V_s dt.$$

62. The process of integration is the inverse of differentiation. Thus, if a differentiated waveform is applied to an integrating circuit, the output is the original waveform. For example, we have seen that a differentiated triangular waveform is a square waveform; consequently, an integrated square waveform is triangular in shape.



(a) PARALLEL TUNED CIRCUIT.



(b) VARIATION OF IMPEDANCE OF A PARALLEL TUNED CIRCUIT WITH FREQUENCY

Fig. 29.—Parallel-tuned circuits

CR circuit as an integrating circuit

63. When a square waveform is applied to a long CR circuit, the capacitor charges and discharges through only a small fraction of the applied voltage during the positive and negative half cycles respectively. Because of this, the rates of charge and discharge remain nearly equal during each half cycle, and the variation of voltage across the capacitor closely approximates to a triangular waveform.

64. Hence, a long CR circuit in which the output is taken across the capacitor is known as an integrating circuit.

Tuned circuits

65. The impedance of a tuned circuit consisting of capacitance in parallel with a series combination of inductance L and resistance R , (fig. 29 (a)), varies with frequency in the way shown by the resonance curve of fig. 29 (b). The circuit has a maximum impedance $R_D = L/CR$ at the resonant frequency f_0 , which is given by

$$f_0 = \frac{1}{2\pi\sqrt{LC}}$$

At this frequency the impedance is resistive, whilst at frequencies above and below f_0 , it is capacitive and inductive respectively.

66. Variation of the series resistance R varies the bandwidth of the circuit, i.e. the range of frequencies for which the impedance is approximately constant. The bandwidth is increased and the impedance at resonance decreased by raising R , i.e. the higher R , the more the circuit is damped (fig. 30). There is always some series resistance present in the form of the HF resistance of the coil.

67. The same damping effect is produced by resistance in parallel with L and C , but for parallel resistance R' , the damping is increased as R' is decreased.

68. If a source of alternating current of variable frequency but of constant amplitude is connected across the circuit, the voltage developed will vary with frequency in exactly the same way as the impedance of the circuit. It will be shown in paragraphs 121 onwards that when an alternating voltage is applied to the grid of a pentode valve, the latter behaves as a generator of current of constant amplitude, if the operating conditions are correctly chosen. Hence the response curve of a tuned

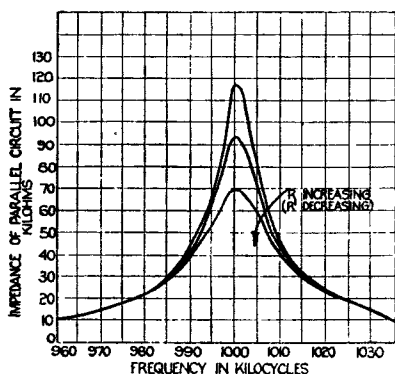
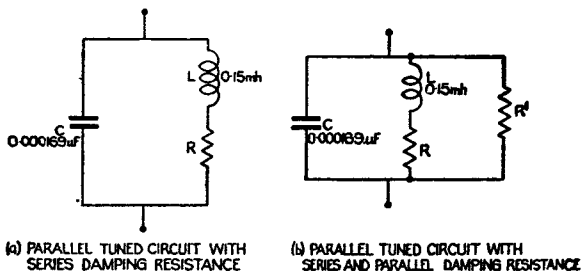


Fig. 30.—Effect of resistance on the impedance frequency curve of a parallel-tuned circuit

circuit connected in the anode circuit of a pentode valve is the same as its frequency: impedance curve.

The Q value of a tuned circuit is defined as the ratio

$$\frac{\text{branch reactance at resonance}}{\text{series resistance}}$$

$$\text{i.e. } Q = \frac{2\pi f_0 L}{R} \quad \text{or} \quad Q = \frac{1}{2\pi f_0 C R}$$

$\therefore$ Increase of R , decreases Q .

In the case of parallel resistance R' ,

$$Q = \frac{\text{parallel resistance}}{\text{branch reactance at resonance}}$$

$$\text{i.e. } Q = \frac{R'}{2\pi f_0 L} \quad \text{or} \quad Q = 2\pi f_0 C R'.$$

69. It can be shown that Q is related to the bandwidth in the following way. Suppose a variable frequency constant current is passed through the circuit, and that f_1 and f_2 are the frequencies on either side of f_0 for which the voltage developed has fallen 3db below that at resonance (fig. 31), then

$$Q = \frac{f_0}{f_2 - f_1}.$$

Over the frequency range f_1 to f_2 , the variation of output is small, so that the quantity $f_2 - f_1$ is commonly called the bandwidth. From the above expression it can be seen that Q is inversely proportional to the bandwidth.

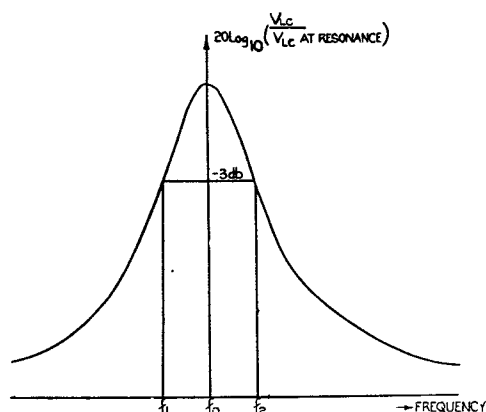
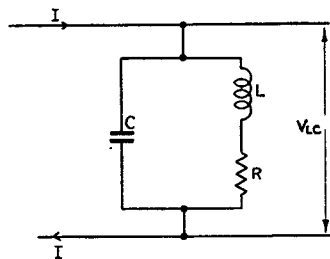


Fig. 31.—Response curve of a parallel-tuned circuit

Response of a tuned circuit to an RF pulse

70. The tuned circuits of the RF and IF amplifiers of radar receivers must be designed to pass pulse-modulated RF without distorting the pulse shape. The response of a tuned circuit to an RF pulse may be assessed either by considering its bandwidth in relation to that required to pass the component frequencies of the pulse or, more directly, by considering the build-up and decay of oscillations in the circuit.

Pulse shape from bandwidth considerations

71. It has been seen that an RF pulse of carrier frequency f_c consists of a range of frequencies extending from $f_c - \frac{1}{4t_1}$ to $f_c + \frac{1}{4t_1}$, where t_1 is the time of rise and fall of the

pulse. In order to pass the pulse without distortion of its shape, a flat response is required over this range of frequencies. This cannot be realized with a single tuned circuit, but may be approximated to by making the bandwidth of the circuit (i.e. the range of frequencies for which the response is not down by more than 3db), equal to the frequency

spread $\frac{1}{2t_1}$ of the pulse.

72. *Example*—To find the bandwidth and Q value of a tuned circuit required to pass RF pulses of carrier frequency $f_c = 45$ Mc/s and

time t_1 of rise and fall = $\frac{1}{10}$ μ sec.

$$\text{Frequency spread } \frac{1}{2t_1} = \frac{1}{2 \times 10^{-7}} \text{ c/s.} \\ = 5 \text{ Mc/s.}$$

Hence required bandwidth = 5 Mc/s.

i.e. from 42.5 to 47.5 Mc/s.

$$Q = \frac{f_c}{\text{bandwidth}} \\ = \frac{45}{5} \\ = 9.$$

Note that the Q value is much lower than that required in normal broadcast technique.

Pulse shape from considerations of build up and decay of oscillations

73. Consider RF current pulses passed through a tuned circuit, e.g. by application of RF voltage pulses to the grid of a pentode in the anode circuit of which the tuned circuit is connected.

Build-up of pulse

74. The RF voltage across the tuned circuit can be shown to build up according to the equation

$$V_{Lc} = V_0 \left(1 - e^{-\frac{R}{2L}t} \right)$$

where

V_{Lc} = peak amplitude of the R.F. voltage
 V_0 = final value of V_{Lc}
= IR_D .
 I = amplitude of supply current
 R_D = dynamic resistance of circuit

R = series resistance

L = inductance.

$$\text{Now } \frac{R}{2L} = \frac{\pi f_c R}{2\pi f_c L} \quad \text{where } f_c \text{ is the resonant frequency of the tuned circuit.}$$

$$= \frac{\pi f_c}{Q}$$

Substituting in the above expression,

$$V_{Lc} = V_0 \left(1 - e^{-\frac{\pi f_c}{Q} t} \right) \\ = V_0 \left(1 - e^{-\frac{t}{Q/\pi f_c}} \right)$$

75. This expression is similar to that for a CR circuit, the quantity $Q/\pi f_c$ being analogous to the time constant of the CR circuit. Thus V_{Lc} attains the value 0.63 V_0 in a time equal to $Q/\pi f_c$, the envelope shape during build-up being as depicted in fig. 32. It is evident that the time required for complete build-up (approximately $5Q/\pi f_c$) is lower the lower the Q of the circuit. Physically, the build-up of oscillations corresponds to the following process. During every oscillation energy is supplied to the tuned circuit from the generator and energy is lost in the tuned circuit due to resistance losses. At the start of the pulse, the energy fed into the circuit from the generator exceeds that lost in the circuit, with the consequence that the oscillation amplitude grows. As the amplitude grows, the energy lost per cycle in the tuned circuit increases. The build-up thus continues until the amplitude is such that as much energy is lost per cycle in the circuit as is supplied per cycle from the generator.

Decay of oscillations

76. When the pulse is removed, oscillations in the tuned circuit die away because energy is dissipated during each oscillation in the resistance of the circuit. It can be shown that the voltage across the tuned circuit decays according to the expression

$$V_{Lc} = V_0 e^{-\frac{R}{2L}t} \\ = V_0 e^{-\frac{t}{Q/\pi f_c}}$$

This expression is again similar to that for a

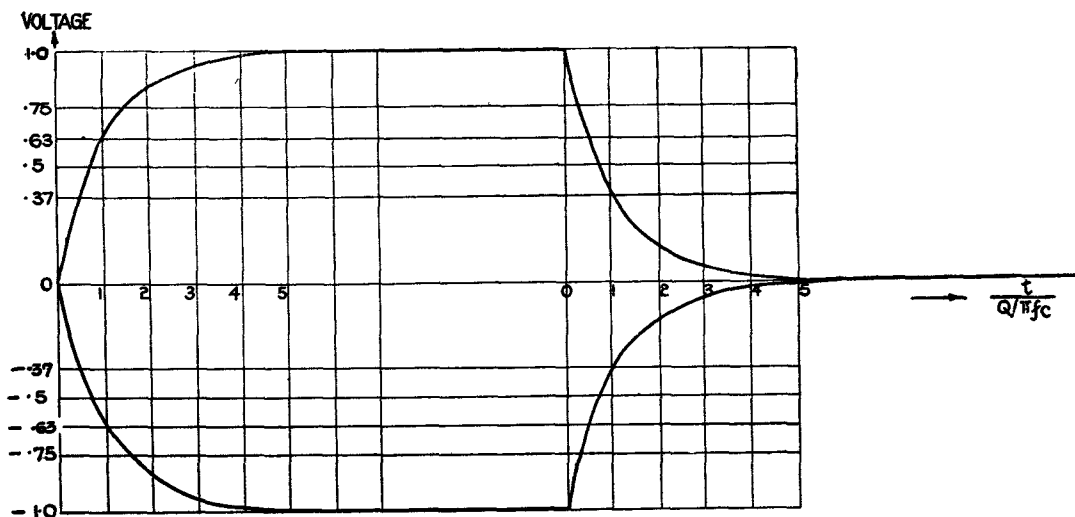


Fig. 32.—RF pulse shape across tuned circuit

CR circuit; the envelope shape is shown in fig. 32.

77. In order to preserve the pulse shape, the build-up and decay of oscillations must be rapid, which, as can be seen from the above expressions, requires that the Q value of the circuit shall be low.

78. *Example.*—To find the Q value of a tuned circuit required to pass R.F. pulses of carrier frequency $f_c = 45$ Mc/s, the time t_1 for rise and fall of pulse being $\frac{1}{10} \mu\text{sec}$.

Time of rise and fall of pulse $\approx \frac{5Q}{\pi f_c}$

$$\therefore \frac{5Q}{\pi f_c} < t_1$$

$$\frac{5Q}{3.142 \times 45 \times 10^6} < 10^{-7}$$

$$Q < \frac{3.142 \times 45 \times 10^6 \times 10^{-7}}{5} \\ < 28.278 \times 10^{-9} \\ < 2.8$$

Note that the Q value determined by this method is very much lower than that calculated from bandwidth considerations. This indicates that a drop of 3db in response at frequencies

$f_c - \frac{1}{4t_1}$ and $f_c + \frac{1}{4t_1}$ is not permissible if the

pulse shape is to conform to the given specifications. The bandwidth method outlined above can thus be considered only as a rough guide for the design of the circuit. From

mathematical considerations it can be shown that the pulse shape across the tuned circuit corresponds to the resultant of two sets of oscillations—an RF pulse of constant amplitude, (A, fig. 33) added to two transients, (B, fig. 33) each consisting of a train of damped oscillations produced in the tuned circuit by the application and removal of the pulse. In the first of the damped trains the oscillations are of opposite phase to those in the RF pulse, the resultant being the difference of the two.

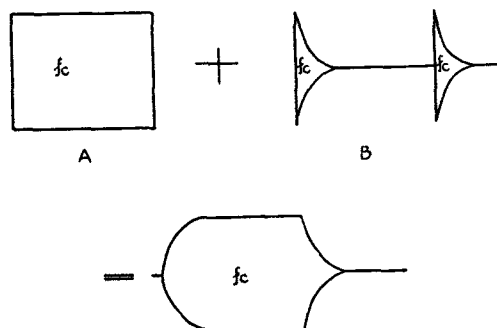


Fig. 33.—Derivation of a pulse shape

Off-tune effects

79. The treatment outlined in the last paragraph makes possible the derivation of the pulse shape obtained when the frequency f_s of the applied pulse differs from the resonant frequency f_c of the tuned circuit. The pulse shape now corresponds to the sum of an RF pulse (A, fig. 34) of frequency f_s and of constant amplitude, and two transients (B, fig. 34) consisting of trains of damped oscillations of frequency f_c . Beats occur at the beginning

of the pulse, decreasing in amplitude as the oscillations of frequency f_c die to zero.

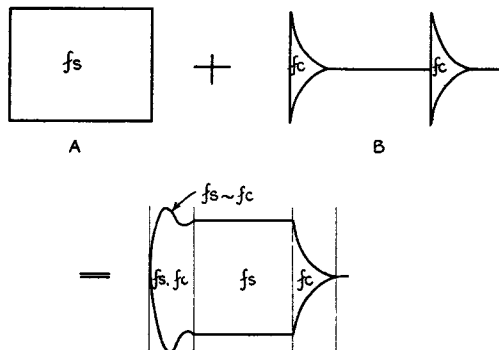


Fig. 34.—Off-tune effects

Ringin circuits

80. When the current flowing through a tuned circuit is suddenly changed, a shocked oscillation or ring is produced. This action occurs in the circuit of fig. 35 when the valve current is alternately switched on and off by the application of a square waveform to the grid.

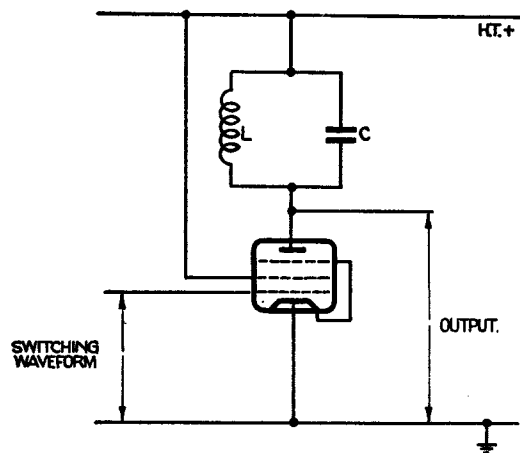


Fig. 35.—Ringin circuit

81. Let us first assume that there is no resistance associated with the tuned circuit and that a pentode with infinite internal resistance R_a is used. Consider that initially the valve is non-conducting and that no current flows in the tuned circuit. At the instant of switching, the whole supply current flows into the capacitor C because current cannot grow instantaneously in the inductance L . As C is charged by the current, its lower plate falls

in potential, a voltage V_{IC} appearing across C and L in parallel which allows growth of current in L (fig. 36). C continues to charge, increasing the (negative) voltage across the tuned circuit and causing the current I_L flowing in the inductance to grow at an increasingly rapid rate. When I_L is equal to the supply current, the current I_c flowing into the capacitor must have fallen to zero. Since a large (negative) voltage exists across the inductance, I_L continues to rise above the

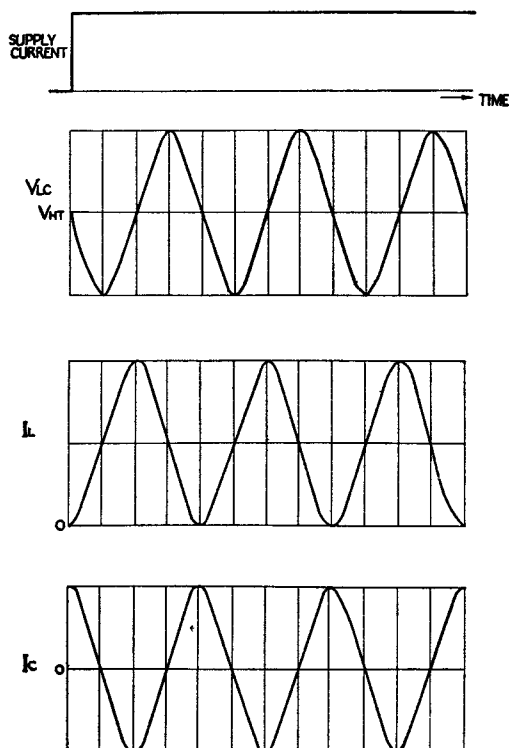


Fig. 36.—Ringin of a resistanceless-tuned circuit

value of the supply current. In order that this may occur, current must flow out of the capacitor which is discharged thereby, the potential of its bottom plate rising. As the discharge of C reduces the (negative) voltage across L , the rate of growth of I_L falls off, reaching zero when C has completely discharged. It can be shown that, at this instant, I_L is equal to twice the value of the supply current (provided there is no resistance in the circuit), and consequently the current flowing out of C is equal to the supply current. The current continues to flow out of C , although it

is discharged, because I_L cannot change suddenly, and in doing so it causes C to develop a voltage of opposite polarity to that to which it was first charged, the potential of the lower plate now rising above V_{HT} . As soon as V_{LC} becomes positive, I_L begins to fall; when it has fallen to the value of the supply current $I_0 = 0$ and the voltage of the lower plate of C is at its maximum positive value. As I_L continues to fall below the value of the supply current, current flows into C , reducing the voltage across it. When V_{LC} reaches zero, $I_L = 0$ and I_c is equal to the supply current, i.e. the circuit has returned to its initial state. This cycle is now repeated, the variation of V_{LC} taking the form of a series of sinusoidal oscillations of constant amplitude.

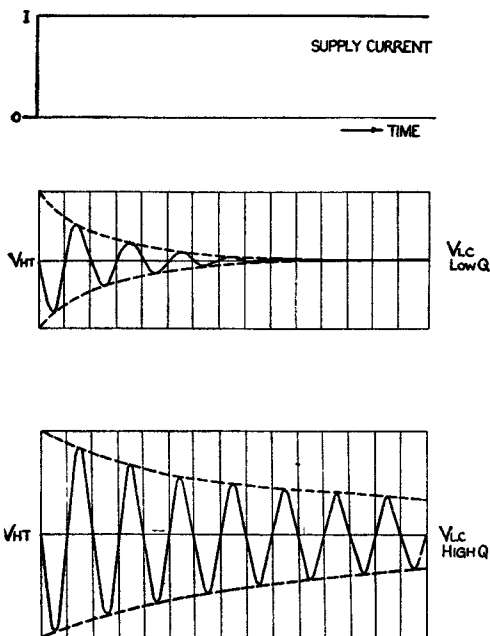


Fig. 37.—Ringing of a damped-tuned circuit

82. If, as in practice, there is resistance associated with the LC circuit, then the oscillations decay in amplitude due to damping by the resistance. Resistance is always present in the form of the HF resistance of the coil, and additional damping may be produced by resistance placed in series or in parallel with the inductance. The rate of decay of oscillations depends on the Q value of the circuit, being slower the higher the Q (fig. 37), as in the case of the decay of an RF pulse. Similar effects are produced when the valve current is switched off, but it can be shown that oscillations start in opposite phase. Fig. 38 shows the oscillations produced across the tuned

circuit of fig. 35 due to the action of the square switching waveform applied to the grid of the pentode.

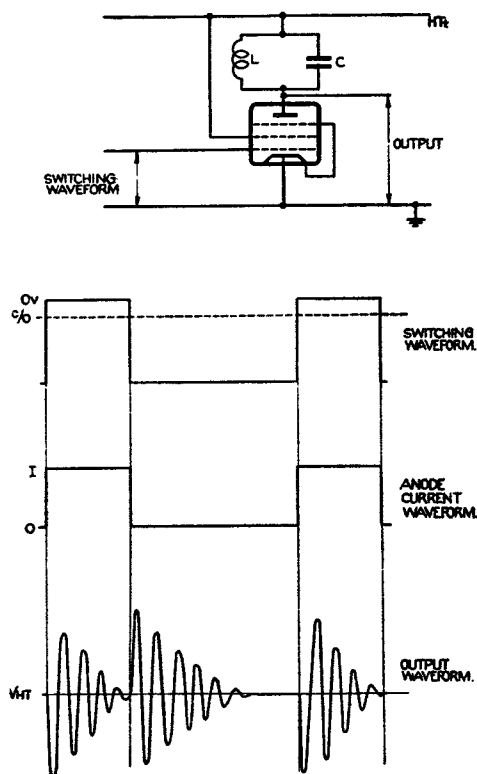


Fig. 38.—Ringing of tuned circuit with pentode switching valve

83. Fig. 39 shows the oscillations produced when a triode is used as the switching valve. The valve is effectively in parallel with the tuned circuit, and since the internal resistance R_a of a triode is low, the damping produced by the triode when conducting is appreciable; thus, the oscillations produced by suddenly switching on current in the valve decay much more rapidly than those caused by switching off the valve current.

84. In order to effectively shock, excite or ring the tuned circuit, the edges of the switching waveform must be steep enough to switch the valve current on or off in a time which is less than the quarter period of the oscillation of the circuit.

85. Ringing circuits are sometimes employed in the calibrator stages which produce voltage markers or pips separated by equal intervals of time to calibrate the time base. The time base triggering waveform is made the switching

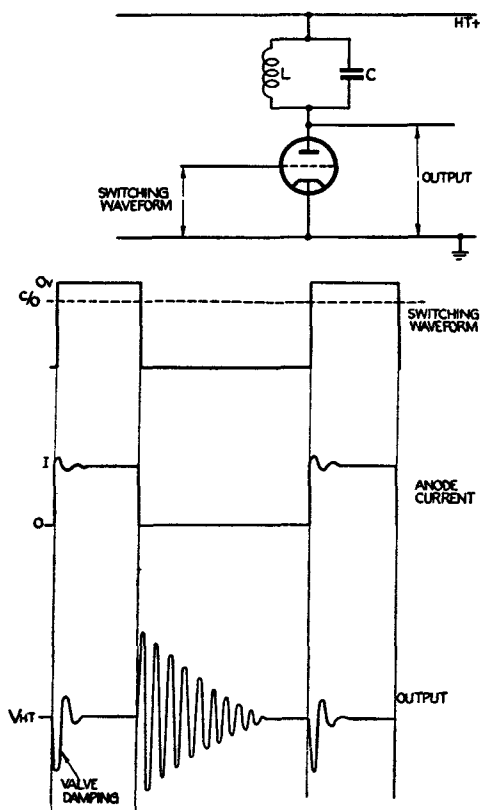


Fig. 39.—Ringing of tuned circuit by triode switching valve

waveform for the ringing valve, and a high Q tuned circuit is used, so that the amplitude of the oscillations does not fall appreciably during the forward stroke of the time base. This sine waveform is squared, and the resulting square waveform applied to a differentiating circuit, which gives a series of pips for application to the Y deflector plates of the display CRT.

86. The ringing circuit can also be used to produce narrow pulses directly. When resistance is placed in parallel with the tuned circuit, the Q value is reduced, and the oscillations decay more rapidly the smaller the resistance, there being a value which damps out the oscillation after the first half-cycle. The waveform across the tuned circuit then consists of a series of positive and negative pips (fig. 40).

DIODE DETECTION

87. In order to produce a deflection or brightening of the time base trace when a signal

is received from a target, such as an aircraft, the RF pulses after amplification in the IF stages of the radar receiver, are detected, the modulation envelope of the pulses being applied to the deflector plates or grid of the display CRT. A diode—which conducts only when its anode voltage is positive with respect to that of its cathode—may be employed to rectify the IF pulses. IF variations on the diode output pulses may be removed by some filter circuit, giving DC pulses as the final output.

88. Detection in radar equipments thus presents a similar problem to that encountered in the broadcast receiver, but more careful design of the filter circuit is required, in order to ensure that efficient filtering of the IF variation is not accompanied by appreciable distortion of the modulation envelope of the pulse.

89. Two possible circuits for a simple diode detector are shown in fig. 41. The operation is the same in both cases, but while the circuit of fig. 41 (a) gives positive pulses, the circuit of fig. 41 (b) gives negative pulses. In the following discussion it is assumed that the first circuit arrangement is used.

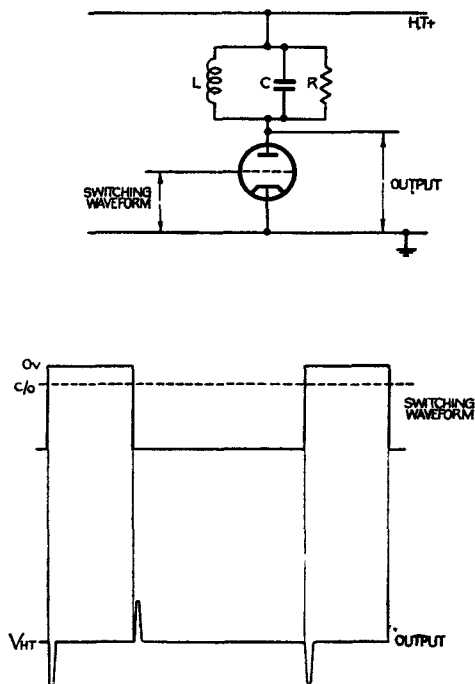
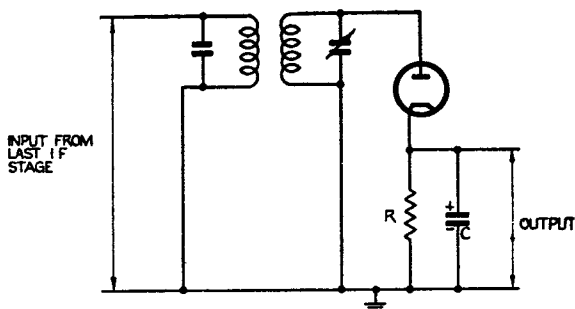
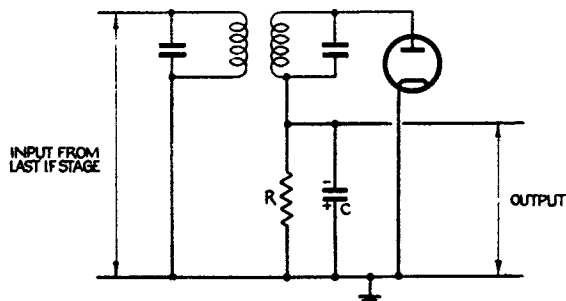


Fig. 40.—Production of narrow pulses from a damped ringing circuit

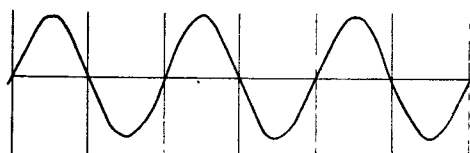


(a) diode detector with positive pulse output



(b) diode detector with negative pulse output

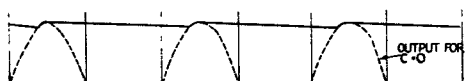
Fig. 41.—Diode detector circuits



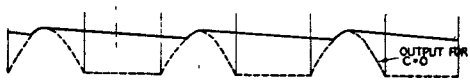
(a) input to detector



(b) output waveform with purely resistive load



(c) output waveform for load with shunt capacitance ($R_D = 0$)



(d) output waveform for load with shunt capacitance ($R_D > 0$)

Fig. 42.—Diode detector waveforms

Sine wave input

90. Consider a continuous sine wave input to the diode, which has a purely resistive load R (i.e. $C = 0$). During the positive half-cycles of the input, the anode-cathode voltage is positive and the diode conducts, but during the negative half-cycles the anode voltage becomes negative and no diode current flows. Thus, the output voltage follows the positive variation of the input, but is zero when the input is negative (fig. 42a and b). If the diode resistance R_D is zero, then no voltage is dropped across the diode when it conducts, and the peak output voltage is equal to the input amplitude. For practical diodes R_D is greater than zero, and so the peak output voltage is reduced by an amount depending on the ratio R to R_D .

The detection efficiency η is defined as the ratio

$$\frac{\text{mean detected voltage}}{\text{peak input voltage}}$$

It may be seen that for a purely resistive load the efficiency of the detector is low, even when a diode with low internal resistance is used.

91. The efficiency can be increased by shunting the load with a capacitor C , which will charge through the diode when this conducts and discharge through R when the diode is cut off. If the diode has no resistance, C charges to the peak value of the input voltage. If C is made so large that it does not discharge appreciably through R before the input voltage again causes the diode to conduct, then the output voltage V_R is practically constant at a value equal to the peak voltage of the input, i.e. $\eta \simeq 1$.

92. Even if the time constant CR is made so long that negligible discharge of C occurs, with practical diodes the mean detected voltage will not be equal to the peak input voltage, unless the load resistance is very large compared with the internal resistance R_D of the diode. The table below indicates the variation of η with R , when CR is large enough to prevent ripple on the output voltage.

$\frac{R}{R_D}$	η
∞	1
100	.9
10	.65
5	.5
0	0

93. Thus two conditions must be satisfied to make the detection efficiency approximate to unity:—

RF pulse input

94. When an RF pulse is applied as input to the diode detector, the latter conducts and C charges up through the diode, the time of rise of the output voltage depending on the time constant CR_D . If CR is long, little RF variation will occur on the output pulse, since C will discharge very little during the portions of the IF cycles which cut off the diode current. However, removal of IF pulse cuts off the diode, and the output voltage can fall to zero only as rapidly as C discharges through R. For long CR the back edge of the output pulse will thus be delayed much more than the front edge (fig. 43b). The pulse shape can be improved by reducing the time constant CR , but this allows more discharge of C when the IF input variations cut off the diode and consequently more IF ripple remains on the output pulse (fig. 43c). Thus a compromise must be made in choosing values of C and R between those which give 100% filtering of the IF from the output and those which preserve the shape of the modulation envelope of the pulse.

95. Since the efficiency is higher the larger the load resistance, it would appear that the CR value should be made low by making R as large as possible and C very small. However there is a lower limit to the value of C which can be used, determined by the capacitance C_D between the electrodes of the diode. C_D and C form a capacitive potential divider, and if C is smaller than C_D , more of the input IF voltage is developed across C than across C_D , i.e. the diode is short circuited by its interelectrode capacitance, and little rectifying action occurs (fig. 43d). Consequently it is necessary for C to be larger than C_D , for appreciable rectification of the IF. For example, if a VR92 or VR78 is used, C_D is approximately 1.6 pF. The load capacitance

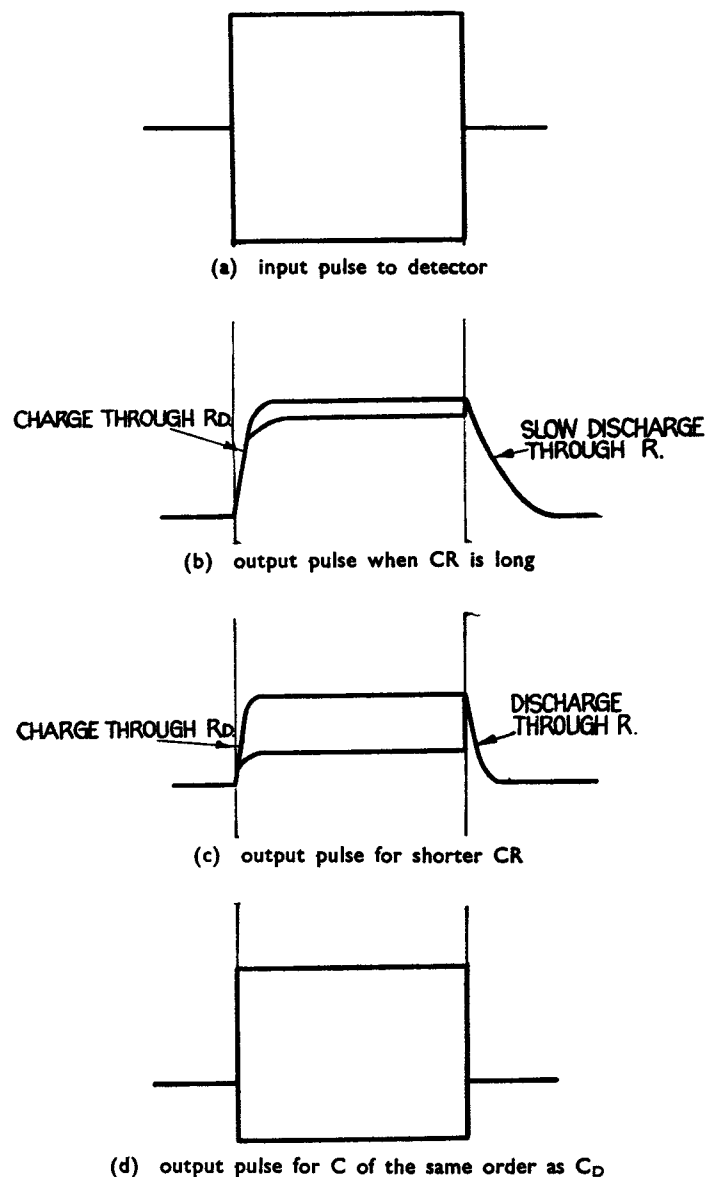


Fig. 43.—Diode detection of RF pulses

(i) The load resistance R must be large in comparison with the diode resistance R_D .

(ii) The time constant CR must be long compared with the period of the IF input.

In broadcast receivers R is of the order 1 M Ω , so that η is high. It is not possible to use such high loads when detecting RF pulses, since the modulation envelope becomes distorted.

must be greater than this value; for 45 Mc/s IF, common values are 5 or 10 pF.

96. The IF ripple remaining on the output pulse is removed by adding a filter across the diode load. The simplest circuit for filtering out the IF variation consists of resistance R_F and capacitance C_F arranged as in fig. 44. Values of R_F and C_F are chosen from the following considerations, the effect of the filter network on the waveform at the cathode of the diode being neglected in the discussion.

The reactance of C_F to the IF, i.e. $\frac{1}{2\pi f C_F}$ must be very much less than R_F , if all the IF variation is to be removed from the output waveform. However, this means that C_F must be large, and consequently the time constant $C_F R_F$ is long. When the voltage levels at the diode cathode are changed rapidly, C_F must charge or discharge through R_F before the output levels can be altered. Thus, to maintain the back edges of the output pulses as steep as in the absence of the filter, the time constant $C_F R_F$ must be at least as small as CR.

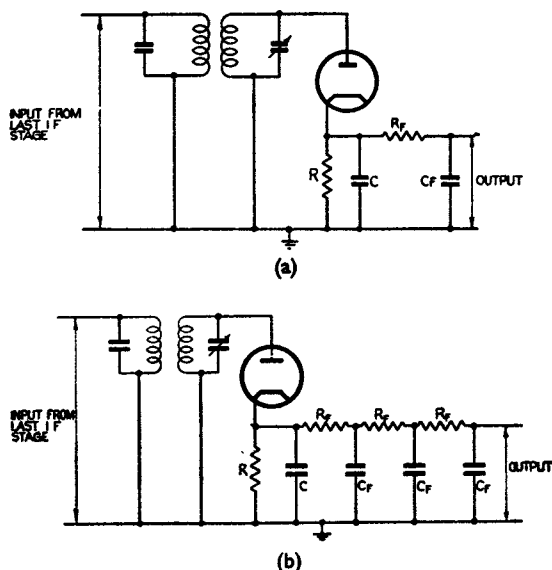


Fig. 44.—Filter outputs for diode detectors

97. With values of C_F and R_F which satisfy the condition,

$$C_F R_F \leq CR$$

the amount of IF variation remaining on the output pulses may still be appreciable. In this case several filter sections all designed to produce minimum distortion of the pulse shape may be used (fig. 44b).

98. Example

A 2 μ sec IF pulse is applied to a diode detector.

What is the largest value of load R which can be used in conjunction with shunt capacitance 5 pF, if the pulse voltage must fall to within 1% of zero in $\frac{1}{5} \times$ pulse width (fig. 45).

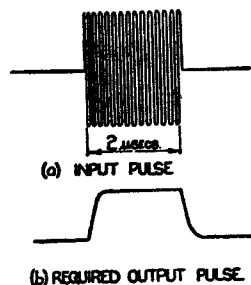


Fig. 45.—DC restorer waveforms (r)

The voltage falls to within 1% of zero in time 5CR.

$$\text{Therefore } 5CR = 0.4 \times 10^{-6} \text{ secs.}$$

$$\begin{aligned} R &= \frac{0.4 \times 10^{-6}}{5 \times 5 \times 10^{-12}} \text{ ohms} \\ &= \frac{400 \times 10^3}{25} \text{ ohms} \\ &= 16K. \end{aligned}$$

Common values of R are 5K and 10K, which give a more rapid decay of voltage.

Assume $R = 10K$, $C = 5pF$. IF = 45 Mc/s. R_F and C_F are added to remove IF ripple.

To preserve pulse shape $R_F C_F \leq CR$.

$$\text{IF } R_F = 10K \quad C_F = 5pF.$$

then for $f = 45 \text{ Mc/s}$

$$\begin{aligned} X_{CF} &= \frac{1}{2\pi f C_F} \\ &= \frac{1}{2 \times 3.142 \times 45 \times 10^6 \times 5 \times 10^{-12}} \text{ ohms} \\ &= \frac{10^5}{3.142 \times 45} \text{ ohms} \\ &= 700 \text{ ohms.} \end{aligned}$$

Thus $X_{CF} \ll R_F$ and most of the IF ripple will be developed across R_F .

DC RESTORATION AND LIMITATION

DC restoration

99. A DC restorer circuit is one which holds either amplitude extreme of a waveform to a given reference voltage level. Such a circuit is sometimes called a "clamping" or "baseline stabilizing" circuit. Fig. 46 shows a square

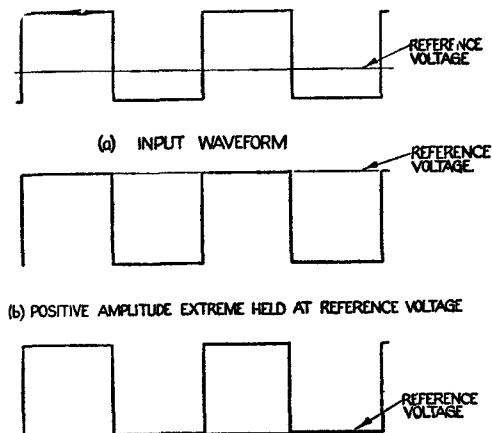


Fig. 46.—DC restorer waveforms (2)

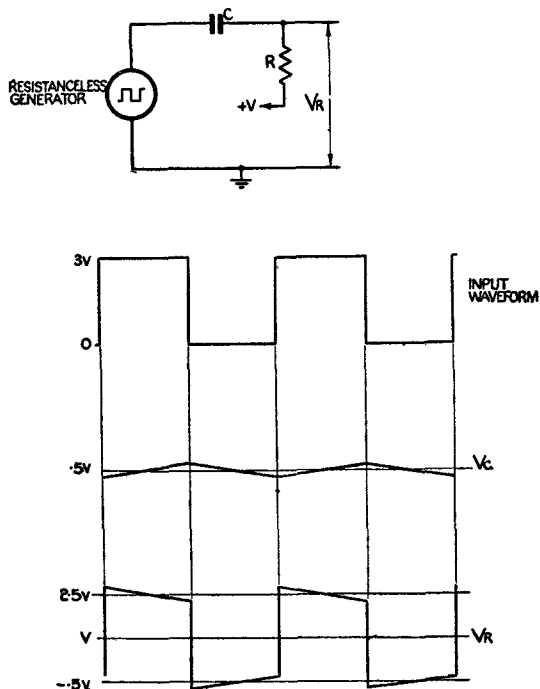


Fig. 47.—Square waveform applied to long CR circuit with R returned to voltage V

waveform input to a DC restorer circuit, and the output when the waveform is allowed to swing only (i) entirely above (ii) entirely below the reference voltage level. The rectifying action of a diode makes it suitable for use as a DC restorer.

Symmetrical square waveform

100. It has been shown (para. 18) that when a square waveform is applied to a long CR circuit, the capacitor C charges to the mean voltage of the applied waveform, while V_R has practically the same shape as the input and its mean level is the voltage to which R is returned (fig. 47). If R is shunted by a diode, the time constants of charge and discharge are unequal and the waveforms across R and C are modified.

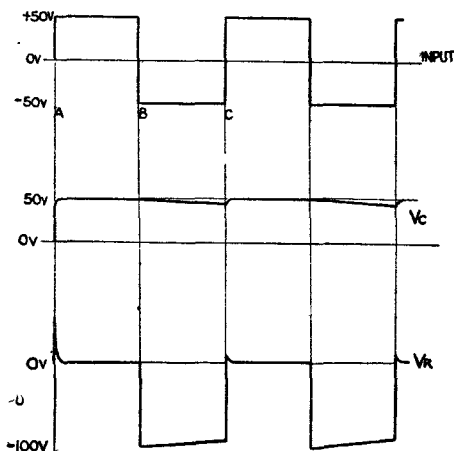
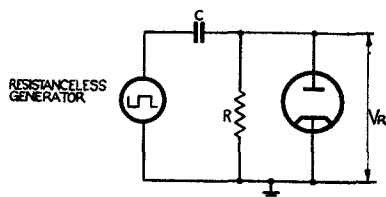


Fig. 48A.—Square waveform applied to negative DC restorer circuit

101. Consider the application from a resistanceless generator of a symmetrical square waveform with voltage levels -50 volts and $+50$ volts to the circuit of fig. 48 (a), C being initially uncharged. At point A, the left-hand plate of C is raised from 0 to $+50$ volts. Since there can be no instantaneous charge of a capacitor, V_R rises by the same amount.

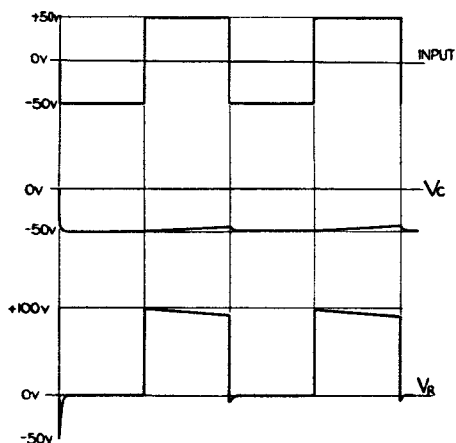
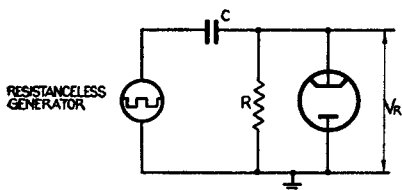


Fig. 48B.—Square waveform applied to positive DC restorer circuit

This makes the diode anode positive with respect to the cathode, and C charges rapidly through the diode to +50 volts, V_R falling to zero at the same rate. These levels are maintained until the input drops to zero at point B. V_R drops from 0 to -100 volts at this instant, since C does not discharge instantaneously. The diode anode is thus taken negative with respect to the cathode, and the diode becomes non-conducting. During the interval BC, C discharges through R, and as V_C falls, V_R rises towards 0 volts. Since CR is long compared with the half-period of the input, little discharge of C occurs.

102. At point C, the input rises from -50 volts to +50 volts, and V_R rises through 100 volts to a slightly positive value. This causes the diode to conduct and charge C rapidly to the peak value of the input. V_R falls to zero at the same rate.

103. In practice the time constant CR is made very long so that V_C varies very little from the peak value of the input voltage, and V_R is a square waveform of the same amplitude as the input, but with its positive level held at zero volts.

104. If the diode connections are reversed, as in the circuit of fig. 48 (b) the action is

similar, but now the diode is made to conduct whenever V_R falls below 0 volts. Consequently, C charges to the negative peak of the input waveform, and the V_R variation is from 0 volts to +100 volts.

Asymmetrical square waveform.

105. The "restoring" action described above always takes place provided the period during which the diode is made conducting is long enough to charge the capacitor to the peak voltage of the input. Thus, an asymmetrical square waveform of period T and with the ratio

$$\frac{\text{duration of positive-going portion}}{\text{duration of negative-going portion}} = \frac{1}{n},$$

applied to the circuit of fig. 47 will be DC

restored as shown in fig. 49 if $\frac{T}{n+1}$ (the dura-

tion of the positive-going portion) is long enough for C to become charged to the peak

positive voltage. If $\frac{T}{n+1}$ is too short, the

positive level of V_R is held at 0 volts, but the amplitude is not as great as that of the input waveform.

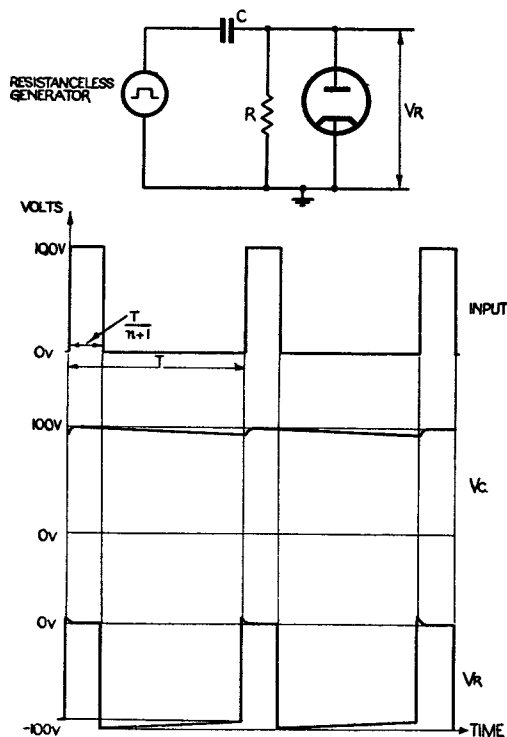


Fig. 49.—DC restoration of asymmetrical square waveform

DC restoration of a sine waveform

106. Consider the application of a sine waveform from a resistanceless generator to a d.c. restoring circuit (fig. 50), the capacitor C being initially uncharged.

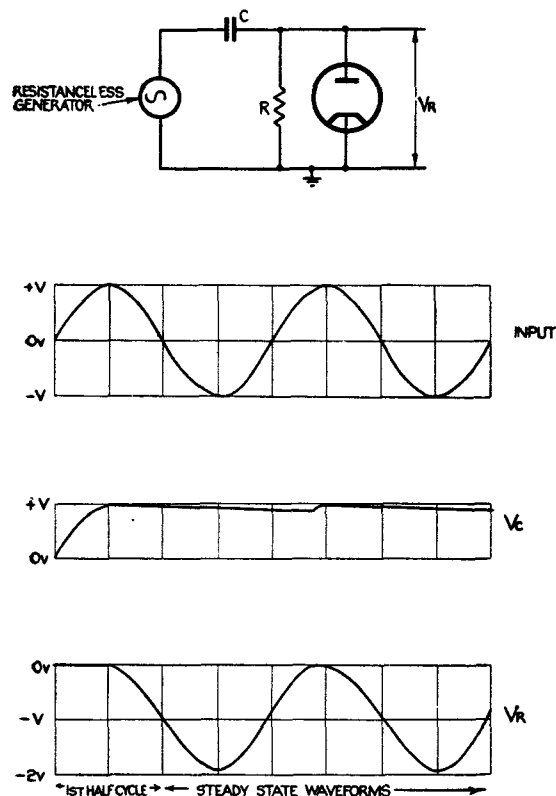


Fig. 50.—DC restoration of a sine waveform

107. During the first quarter-cycle, the anode of the diode is taken positive with respect to the cathode, and C charges to the peak input voltage V through the diode as rapidly as the input voltage rises. Consequently, no voltage appears across R during this period.

108. At the beginning of the second quarter-cycle, the voltage of the diode anode falls below that of the cathode, and C begins to discharge slowly through R . C continues to discharge through R for the rest of the cycle, while V_R goes negative at almost the same rate as the input voltage falls, so that at every instant $V_C + V_R$ is equal to the input. The maximum negative value of V_R (at the end of the third quarter-cycle) is very nearly equal to $-2V$ if CR is very long compared with the period of the sine waveform.

109. C continues to discharge in the fifth quarter-cycle until the input voltage rises to the value of V_C ; at this instant $V_R = 0$. Further rise of the input voltage brings the diode into conduction and C charges to the positive peak voltage V through the diode, V_R remaining at zero.

110. When the input falls below $+V$ again, the diode cuts off, and the whole action from the beginning of the second quarter-cycle then repeats itself in successive cycles. In the steady state, V_C is more or less constant at the peak voltage V , while V_R is a sine wave similar to the input, but slightly flattened at the positive peak which is held at zero volts by the d.c. restoring action.

DC restoration to any potential

111. In the above discussion it has been assumed that one electrode of the diode is returned to earth, so that the reference level in each case has been zero. This is not a necessary condition for d.c. restoration. If a symmetrical square waveform is applied to a

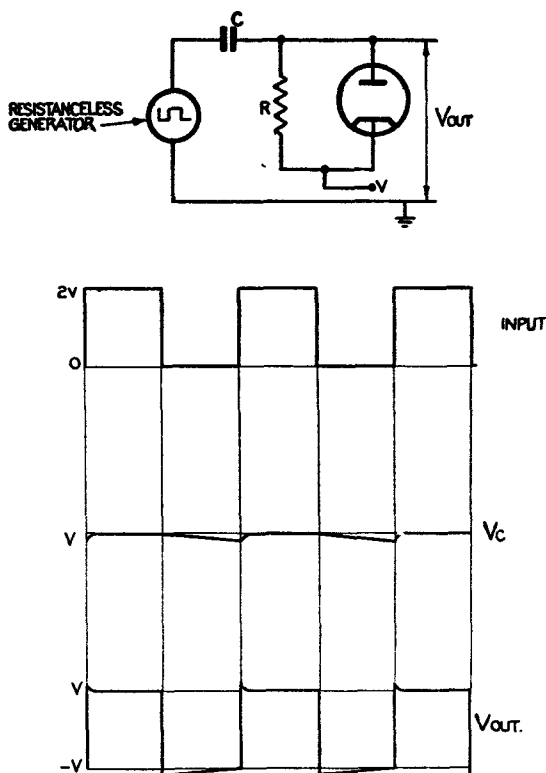


Fig. 51A.—Negative DC restoration to potential V

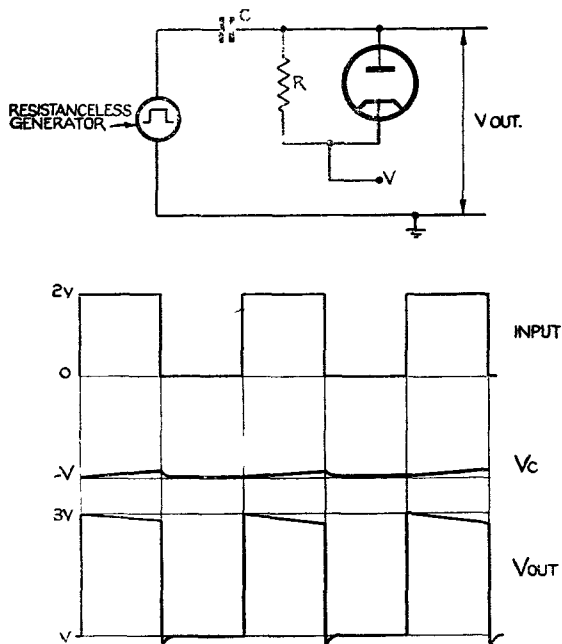


Fig. 51B.—Positive DC restoration to potential V

long CR circuit in which R is returned to voltage V , then the output square waveform swings equally above and below the level V . When R is shunted by a diode, connected as in fig. 51 (a) the V_R variation above V causes the diode to conduct and charge C rapidly to the peak of the input voltage. Thus the positive level of the waveform across R becomes equal to V .

112. Similarly, if R and the anode of the diode are held at voltage V , the diode conducts, V_R falls below V and d.c. restoration occurs, the negative level of V_R now being held at the reference level V (fig. 51 (b)).

DC restoration in the grid circuit of a valve

113. DC restoration may occur at the grid of a triode or pentode valve if an input signal is applied via a CR coupling (fig. 52). If the input signal is large enough to take the grid positive with respect to the cathode, grid current flows and the reactance of the grid cathode space becomes very low. C_G charges rapidly to the peak input voltage, and V_{gk} , the grid cathode voltage, falls to zero at the same rate. When the input voltage swing takes the grid negative, grid current stops flowing and the grid cathode space becomes a high reactance, and C_G now discharges slowly through the grid leak R_g . Thus, the grid

waveform has its positive level clamped to earth, with no bias, and to the potential of the cathode if a bias voltage is used.

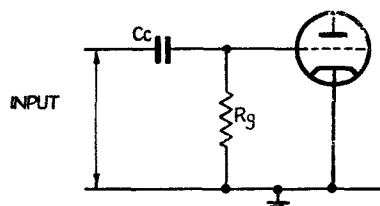


Fig. 52.—Grid clamping circuit

Limitation

114. A "limiter" circuit is one which removes the part of an input waveform which lies above or below a reference voltage level. Since a diode conducts only when its anode is positive with respect to the cathode, it is often used as a limiter. There are numerous limiter circuits in common use but only parallel-diode limiting will be discussed here. A typical circuit arrangement is shown in fig. 53 (a). The diode is connected in parallel with the load, which is assumed to be of very high impedance, so that the current flowing through it is negligible. In series with the diode is a resistor R , which is very much larger than the resistance of the diode when conducting. We shall consider the application to the circuit of a symmetrical triangular waveform from a generator which has zero internal resistance, the voltage extremes of the waveform being $+V$ and $-V$ (fig. 53 (b)). When the input voltage is negative, the diode does not conduct, and so acts as an open circuit. Consequently, no voltage is developed across R , and the output follows the input voltage. When the input swings above earth potential, the anode is

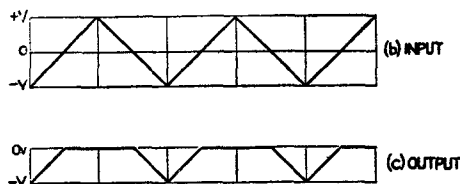
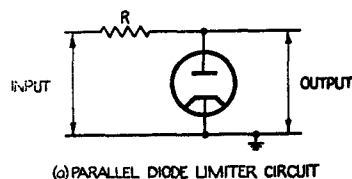


Fig. 53.—Positive limiting to earth potential by a parallel diode limiter

taken positive with respect to the cathode and the diode begins to conduct, its resistance falling from an infinitely high value to a value of the order of 500 to 1000 ohms. During the positive half-cycle, therefore, potential division occurs between R and the low resistance of the diode, and provided R is very much larger than this resistance, the output voltage remains at zero (fig. 53 (c)).

115. If the diode connections are reversed as in fig. 54 (a), so that the anode is held at earth potential, then the diode conducts during the negative half-cycles only. The diode current flows through the resistance R , developing a voltage across it appreciably equal to the input voltage. Thus the output voltage follows the input during the positive half-cycles but remains at zero during the negative half-cycles (fig. 54 (c)).

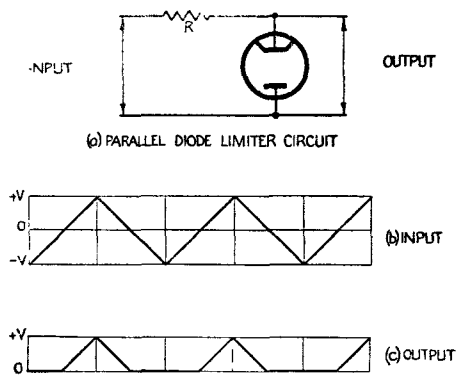


Fig. 54.—Negative limiting to earth potential

116. An input waveform can be limited to any positive or negative value by holding the appropriate diode electrode at the required potential. If the cathode is returned to a positive potential, the diode does not conduct until the input causes the anode voltage to exceed this value. Similarly, if the cathode is held at a negative voltage, the diode conducts when the anode is positive and when it is at negative voltages which are less negative than that of the cathode.

117. Examples:—

(i) In fig. 55 (a), the cathode is connected to a battery of +20 volts. The part of the input waveform which takes the anode above +20 volts causes the diode to conduct and become a very low resistance. During this part of the input waveform, the anode or out-

put voltage is thus equal to +20 volts. For the remainder of the cycle, when the input voltage is less than 20 volts, the diode does not conduct, and the output equals the input.

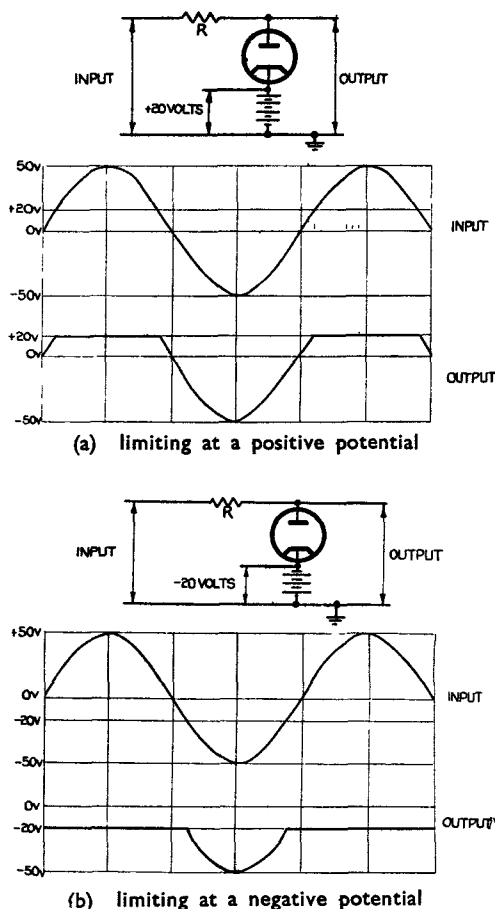
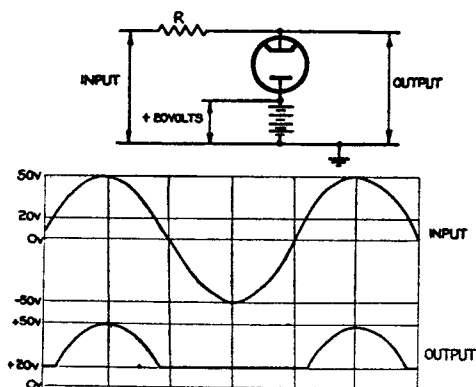


Fig. 55.—Positive limiting by a parallel diode circuit

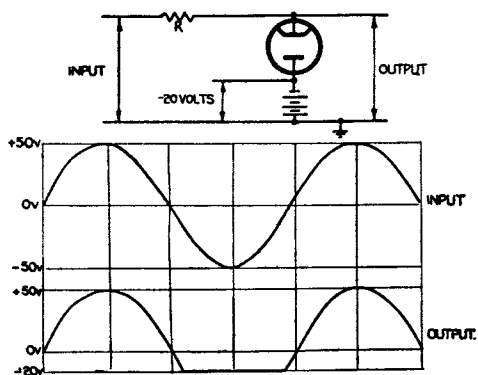
(ii) When the cathode is held at -20 volts as in the circuit of fig. 55 (b), the diode conducts when the input voltage is more positive than -20 volts, and the output voltage remains at this level, until the input becomes more negative than -20 volts. The diode is then cut off, and the output follows the input waveform.

If the diode connections are reversed, then the waveforms are as shown in fig. 56 (a) and (b).

118. In practical circuits the waveform generator has some internal resistance, e.g. if the generator takes the form of an amplifier with



(a) limiting at a positive potential

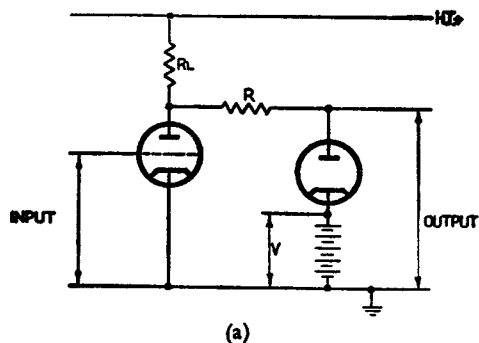


(b) limiting at a negative potential

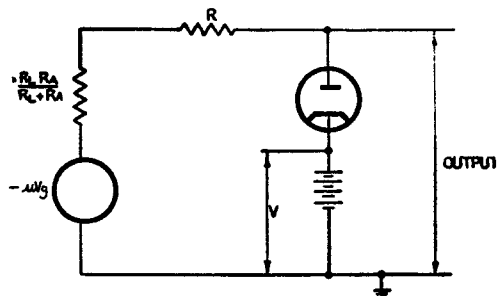
Fig. 56.—Negative limiting by a parallel diode circuit

a resistive load R_L (fig. 57 (a)), then its effective resistance R_I , is equal to R_L in parallel with the slope resistance R_a of the valve (fig. 57 (b)). Thus potential division occurs between $(R + R_I)$ and the diode resistance when the diode conducts. When the diode is non-conducting, no current flows through R and the output voltage is equal to the anode voltage of the amplifier. It would thus appear that an efficient limiter circuit could be obtained by connecting the diode electrode directly to the anode of the amplifier, but for efficient limiting to the voltage at which the other diode electrode is held, R_I must be very much larger than the diode resistance, otherwise an appreciable voltage will be produced across the diode, causing the maximum current rating of the diode to be exceeded.

119. Limiters are useful in squaring off the extremities of waveforms. A pulse waveform in which the top of the pulse is not level can be squared in this way (fig. 58), and its amplitude adjusted to a required value. Another

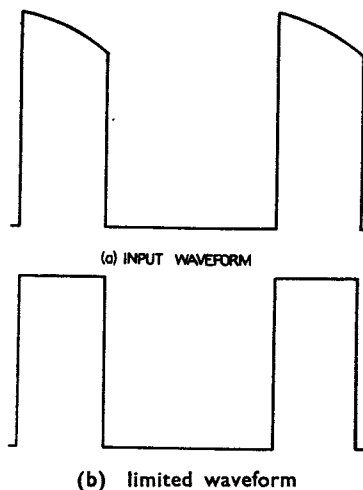


(a)



(b)

Fig. 57.—Amplifier and parallel diode limiter



(b) limited waveform

Fig. 58.—Limiter used for squaring

use of a limiter circuit is to remove either the positive or negative peaks from a differentiated square waveform (fig. 59).

Double-diode limiting

120. A square waveform can be obtained by limiting a sine waveform by a limiter circuit which has two diodes in parallel. In the circuit of fig. 60, V_1 conducts whenever the input voltage rises above $+V$, thus limiting

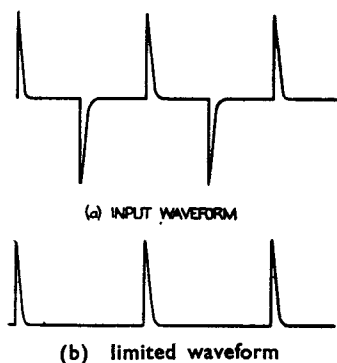


Fig. 59.—Removal of positive or negative peaks from differentiated square waveform

the positive half-cycle to the potential $+V$. Similarly, V_2 limits the negative half-cycle to the potential $-V$.

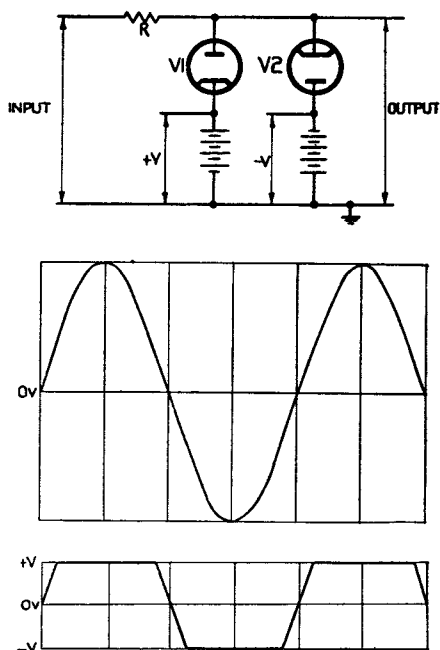


Fig. 60.—Double diode limiting

VALVE THEORY

121. Before discussing typical circuits used in radar equipments, the characteristic curves and valve constants for triodes and pentodes will be briefly revised.

122. The characteristic curves for a typical triode valve are shown in fig. 61. Fig. 61 (a)

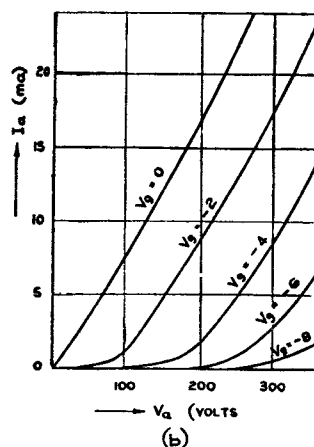
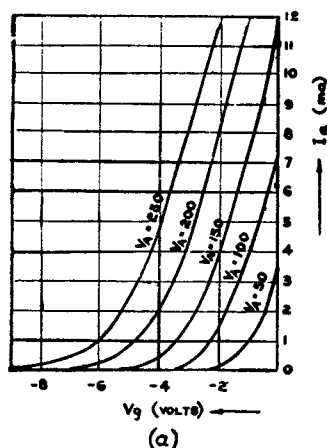


Fig. 61.—Characteristic curves of a triode (MH4)

shows the variation of anode current I_a with grid-cathode voltage V_g for various values of anode-cathode voltage V_a , while in fig. 61 (b), I_a is plotted against V_a for different values of V_g .

123. The pentode characteristics are shown in fig. 62. Fig. 62 (a) gives the $I_a:V_g$ curves for several values of V_a , the screen-cathode voltage V_{sg} being maintained constant; fig. 62 (b) shows the $I_a:V_g$ curves for constant V_a and different values of V_{sg} ; fig. 62 (c) shows the $I_a:V_a$ curves for several values of V_g, V_{sg}

being kept constant. It can be seen from fig. 62(c) that change of V_a has little effect on I_a except for low values of V_a where the $I_a:V_a$ curve rises steeply. Fig. 62(a) and (b) show that V_{bg} has much more control than V_a on I_a .

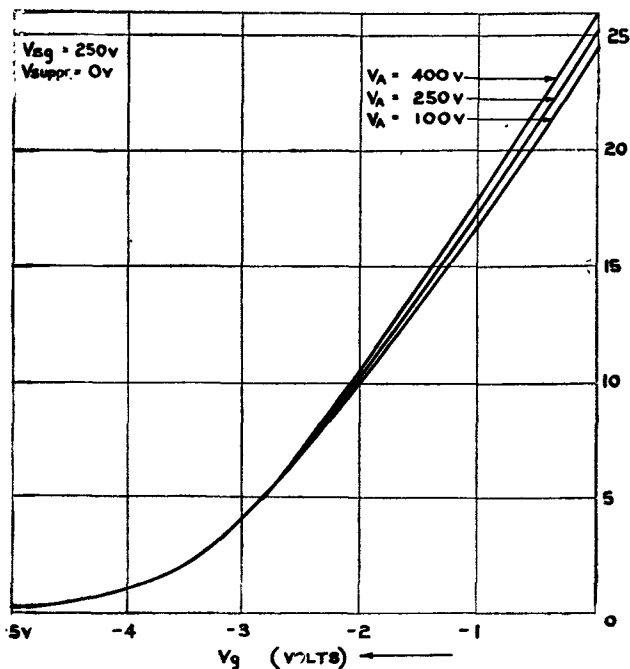
Valve "constants"

124. The amplification factor μ is defined as follows:—

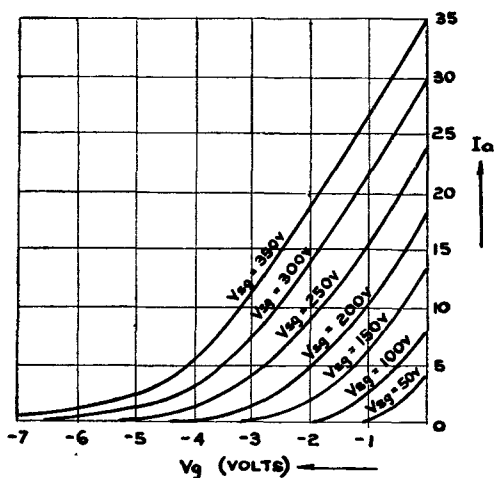
$$\mu = \frac{\text{change of } V_a \text{ which produces a given change in } I_a}{\text{change of } V_g \text{ which produces the same } I_a \text{ change}}$$

$$= \left(\frac{dV_a}{dV_g} \right) I_a$$

= ratio of effectiveness of grid and anode voltages in controlling I_a .



(a)



(b)

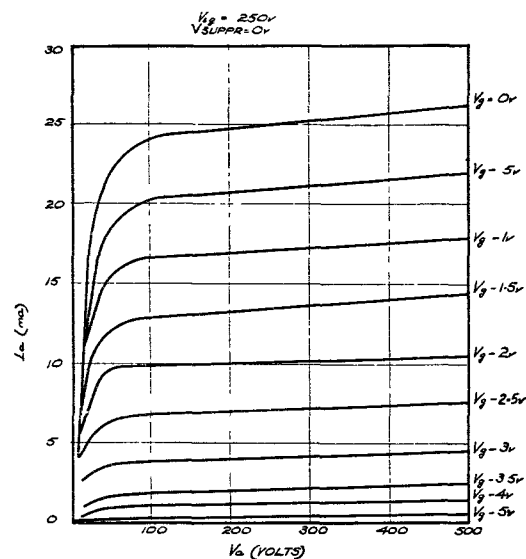
125. Because of dissymmetry of valve construction, μ is not a constant, but varies with the operating conditions of the valve. The variation of μ with operating conditions, and the variation from valve to valve can be estimated by examining the valve characteristics. It can be seen that

(i) the smaller the horizontal separation of the $I_a:V_g$ curves for a given V_a difference

(ii) the greater the horizontal separation of the $I_a:V_a$ characteristics for a constant V_g difference

the larger is μ .

126. Thus, over the steeply-rising portion of the $I_a:V_a$ curves for the pentode μ tends to zero, but for higher values of V_a where the curve is almost a horizontal line, μ is extremely high. This is due to the fact that the anode is screened from the cathode so that increase of V_a does not alter the value of I_a . The characteristics also show that μ is very much



(c)

Fig. 62.—Characteristic curves of a pentode (EF50)

higher for a pentode than a triode. Triodes have μ values ranging from 3 to 100; pentodes from 250–2000 except for very low V_a where μ tends to zero.

127. The anode or slope resistance R_a represents the resistance offered by the anode circuit to a small increase in I_a , i.e.

$$R_a = \frac{\text{change in } V_a}{\text{change in } I_a \text{ produced}} \quad (V_g \text{ constant})$$

$$= \left(\frac{dV_a}{dI_a} \right) V_g$$

Like μ , R_a is not a constant. From the characteristics we see that

(i) the smaller the vertical separation of the $I_a:V_g$ curves for a constant V_a difference

(ii) the smaller the slope of the $I_a:V_a$ characteristics

the larger is R_a .

128. For triodes R_a lies between 1K and 100K. For pentodes R_a lies between 50K (for power pentodes) and values of the order of 1M, except for low V_a where R_a is low.

129. The mutual conductance g_m is defined as

$$g_m = \frac{\text{change of } I_a}{\text{change of } V_g \text{ producing it}} \quad (V_a \text{ constant})$$

$$= \left(\frac{dI_a}{dV_g} \right) V_a$$

g_m also varies over the characteristics, being large

(i) the greater the slope of the $I_a:V_g$ characteristics

(ii) the greater the vertical separation of the $I_a:V_a$ curves for a constant V_g difference

g_m has the values between 1 and 15 ma/volt.

130. The three valve constants are related by the equation

$$\mu = g_m R_a$$

$$\text{for } g_m R_a = \left(\frac{dI_a}{dV_g} \right) V_a \times \left(\frac{dV_a}{dI_a} \right) V_g$$

$$= \left(\frac{dV_a}{dV_g} \right) I_a$$

$$= \mu$$

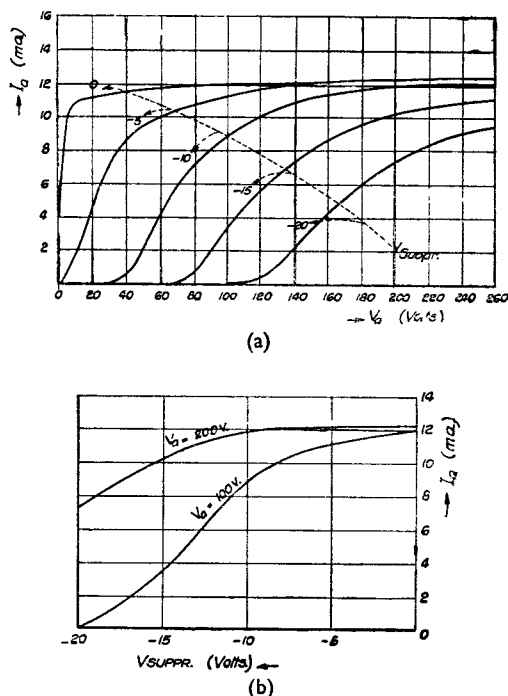


Fig. 63.—Suppressor grid characteristics for pentode

131. It is quite common practice in radar to apply modulating waveforms to the suppressor grid of pentode valves, and so the suppressor characteristics of a typical pentode are shown in fig. 63.

132. Fig. 63 (a) shows the $I_a:V_a$ curves for various values of V_{suppr} , while V_g and V_{sg} are kept constant. When the suppressor is held at the same voltage as the cathode, electrons moving between the screen and suppressor are slowed down by the retarding field, but except at very low anode voltages, the electrons pass through the suppressor grid and reach the anode without stopping. If the suppressor is made negative, the anode must be made more positive before it can produce a sufficiently strong electrostatic field to draw off the electrons as rapidly as they arrive at the screen grid side of the suppressor. Thus electrons collect and form a space charge or "virtual cathode" near the suppressor. This space charge, the suppressor grid and the anode act in the same way as the cathode, control grid and anode of a triode valve, and so for negative values of the suppressor, the curves are like

the $I_a \cdot V_a$ curves of a triode, the R_a value being comparatively low. Fig. 63 (b) shows the variation of I_a with V_{suppr} when V_a , V_g and V_{sg} are kept constant.

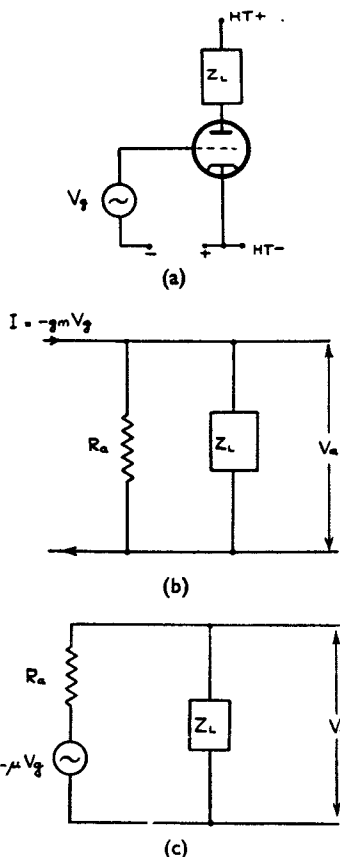


Fig. 64.—Equivalent circuit of a voltage amplifier

The equivalent circuit of a voltage amplifier

133. In the circuit of fig. 64 (a) a load Z_L is connected between the anode of a valve and the HT supply, the valve being biased so that the control grid is always slightly negative with respect to the cathode. The valve may be either a triode or pentode. The grid voltage is varied by applying a signal voltage V_g . The consequent variation of anode voltage may be determined as follows:—

134. If a change V_g in the signal voltage makes the grid more positive, then the anode current is increased. This increased current flowing through Z_L , causes a change in anode voltage. The change I_a in anode current is determined by both V_g and the anode voltage change V_a .

$$I_a = g_m V_g + \frac{V_a}{R_a}$$

Now, since I_a is an increase, there will be a larger voltage drop across Z_L and the anode voltage falls.

$$\text{i.e. } V_a = -I_a Z_L$$

$$= -Z_L \left(g_m V_g + \frac{V_a}{R_a} \right)$$

$$\therefore V_a \left(1 + \frac{Z_L}{R_a} \right) = -g_m V_g Z_L$$

$$V_a = - \frac{R_a Z_L}{R_a + Z_L} \cdot g_m V_g \quad \dots\dots (A)$$

or if μ is substituted for the product $g_m R_a$, we get

$$V_a = - \frac{Z_L}{R_a + Z_L} \cdot \mu V_g \quad \dots\dots (B)$$

135. Expressions (A) and (B) for the change in anode voltage indicate that the voltage amplifier can be represented by the equivalent circuits of fig. 64. Expression A indicates that the valve amplifier with grid voltage fluctuation V_g may be regarded as generating a current $-g_m V_g$, which flows through the anode resistance R_a of the valve in parallel with the load impedance Z_L (fig. 64 (b)). This is known as the constant current generator form of the equivalent amplifier circuit. An alternative equivalent circuit—that of the constant voltage generator—can be derived from expression B. This shows that the anode current which flows through the load when a voltage change V_g is applied to the grid, can be calculated by replacing the anode cathode circuit of the valve by a generator of voltage $-\mu V_g$ in series with the anode resistance R_a of the valve (fig. 64 (c)).

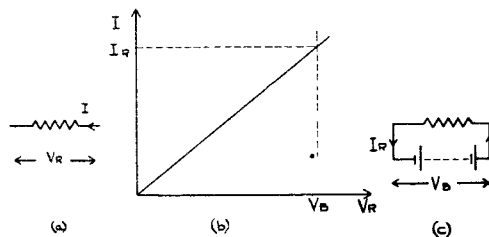


Fig. 65.—Current-voltage graph for single resistor

136. These equivalent circuits are useful to simplify circuit problems. It is convenient to use the constant current generator form when the anode resistance R_a is very much higher than the load, as is generally the case with pentode and tetrode valves. Expression A then approximates to the form

$$V_a = - \frac{R_a Z_L}{R_a} \cdot g_m V_g$$

$$= - Z_L \cdot g_m V_g.$$

and the amplification A from grid to anode is given by

$$A = \frac{V_a}{V_g} = - g_m Z_L$$

This expression suggests that A is directly proportional to the load Z_L . However, g_m decreases for very high values of Z_L —as will be seen from a study of load lines—so that A tends to reach an upper limit.

137. The equivalent voltage generator circuit is most useful when R_a is of the same order as or less than the load impedance, which is commonly the case for the triode valve amplifiers. If the load impedance is made very much larger than R_a , then expression B becomes

$$V_a \rightarrow - \frac{Z_L}{Z_L} \cdot \mu V_g.$$

$$\rightarrow - \mu V_g.$$

and the amplification $A = \frac{V_a}{V_g}$

$$= - \frac{\mu Z_L}{R_a + Z_L}$$

$$\rightarrow - \mu, \text{ where } Z_L \gg R_a.$$

Thus as Z_L is increased, A increases, but since the value of R_a increases for very high anode loads, the rate of increase of amplification then falls off. This may be seen by a study of the load lines, a load line being a graph of anode voltage against anode current, superimposed on the $I_a:V_a$ characteristics of the valve.

Load lines

138. Consider a current I flowing through resistance R. Then the voltage V_R developed across the resistor is equal to IR , and increase of I causes a proportional increase of V_R . Thus a graph of I plotted against V_R takes the

form of a straight line (fig. 65 (b)), the slope of the line being smaller the greater the resistance. If now a battery of voltage V_B is connected across the resistor, the current which flows through it can be determined by reference to the graph (fig. 65 (b) and (c)).

139. Now consider a battery of voltage V_B connected across two resistors R_1 and R_2 in series (fig. 66 (a)). The current: voltage curve or load line for R_1 will again be a straight line passing through the origin of the axes. To determine the curve for R_2 , assume first of all that the circuit is broken at the junction point A and that current is taken from the battery through R_2 . If no current is taken, the voltage of point A will be equal to the battery voltage V_B . If a current I is taken, a voltage IR_2 is developed across the resistor and the voltage of point A falls to $V_B - IR_2$. The larger the current I, the larger IR_2 becomes, and the lower the voltage of point A. Thus the load line or current: voltage curve for R_2

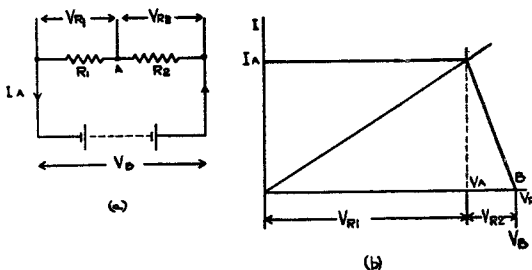


Fig. 66.—Current-voltage graph for two-series resistors

is a line passing through the point B ($V_R = V_B, I = 0$) the slope being smaller the larger the value of R_2 . The voltage of the point A, (V_A) and the current I_A flowing in the circuit of fig. 66 (a) are determined by dropping perpendiculars to the axes from the point of intersection of the two load lines of fig. 66 (b); if a current greater or less than I_A is assumed to be flowing then it can be seen from fig. 66 (b) that the sum of the voltages V_{R1} and V_{R2} developed by this current flowing through the resistors will be greater or less respectively than the voltage V_B applied across them. Thus, the only possible current which can flow is that given by the point of intersection of the load lines.

140. When the resistor R_1 is replaced by a valve, we have the circuit of a voltage amplifier

(fig. 67 (a)). Since the $I_a:V_a$ curve for a valve varies with the control grid voltage V_g , the single load line for R_1 has to be replaced by a family of characteristic curves for the valve for different values of V_g . Then, for a particular value of V_g , the anode current I_a flowing and the anode voltage V_a are given by intersection of the $I_a:V_a$ curve for this V_g value with the load line for the anode load resistance R_2 .

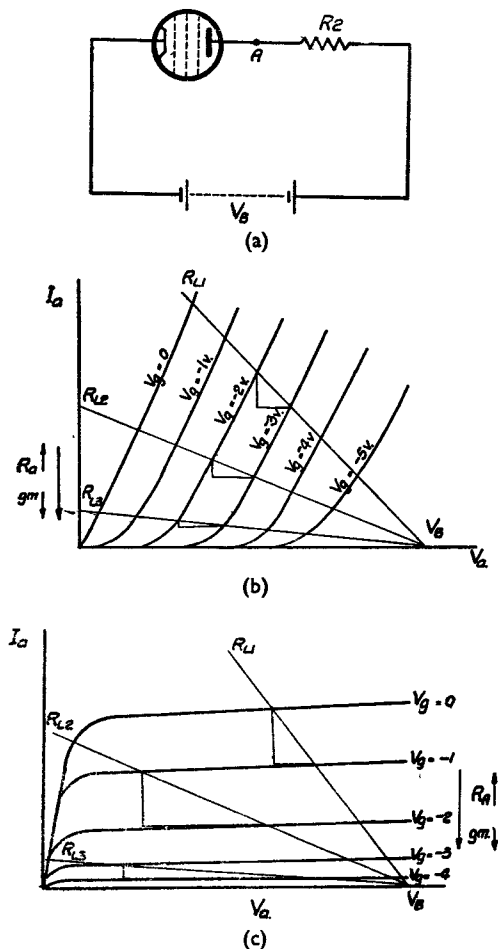


Fig. 67.—Load lines

141. In fig. 67 (b), the load line for resistor R_1 is replaced by the $I_a:V_a$ curves of a triode and in fig. 67 (c) by the $I_a:V_a$ characteristics for a pentode. Load lines for resistive loads $R_{L1} < R_{L2} < R_{L3}$ (corresponding to resistor R_2 above) are superimposed on each diagram.

142. At the points where the load line for R_{L1} intersects the characteristics of fig. 67 (b) their slope is high but it becomes smaller for the R_{L2} load line, and even more so for the

R_{L3} line. Since R_a increases with decreasing slope of the $I_a:V_a$ characteristics, the R_a corresponding to the load R_{L1} is smaller than that for R_{L2} , which is smaller than that for R_{L3} . Similarly, it can be seen from fig. 67 (c) that for a pentode g_m decreases with increase of R_L . Thus for both pentodes and triodes the amplification does increase with the load, but for large load resistances, the rate of increase of amplification is much lower than that of the load resistance.

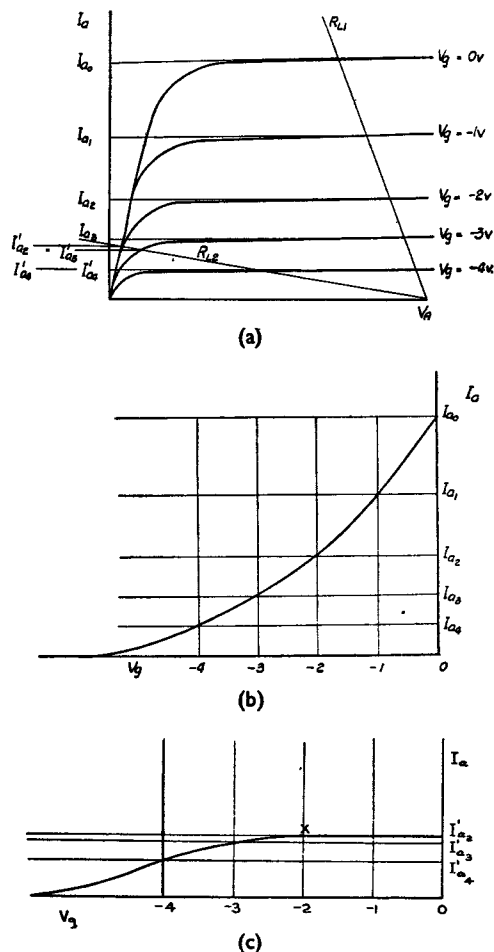


Fig. 68.—Anode bottoming and dynamic characteristic for anode load

143. With a pentode it is possible to amplify an input signal about 200 times, but with a triode the maximum amplification obtainable is of the order of 50.

Anode bottoming

144. The $I_a:V_a$ characteristics of a pentode are sketched in fig. 68 with load lines for two

resistive anode loads $R_{L1} < R_{L2}$ superimposed. Consider first the smaller load R_{L1} . As V_g is raised (i.e. made less negative), I_a increases and V_a falls. The screen current I_{sg} also increases slightly. For $V_g = -4, -3, -2, -1, 0$ volts respectively, the corresponding anode current values are $I_{a4}, I_{a3}, I_{a2}, I_{a1}, I_{a0}$. But if the load R_{L2} is used, then for $V_g = -4, -3, -2, -1, 0$ volts, the corresponding values of I_a are $I'_{a4}, I'_{a3}, I'_{a2}, I'_{a1}, I'_{a0}$, i.e. for all values of V_g above -2 volts, the intersection of the load line with the $I_a:V_a$ curves is common, and I_a remains constant. However, as V_g is raised above -2 volts, the total current in the valve continues to increase, and since I_a cannot rise, the total increase in current is taken by the screen. For this value of R_{L2} , the pentode is said to "bottom" below -2 volts. It will be seen later that anode bottoming in pentodes is of great importance.

Dynamic characteristics

145. The dynamic characteristic for a valve graphs the anode current flowing in the valve when it has an anode load in circuit, against the grid-cathode voltage. It can be derived from the static $I_a:V_a$ characteristics as follows. Consider a pentode amplifier with the static $I_a:V_a$ characteristics shown in fig. 68 (a) and anode load R_{L1} . The intersections of the load line for R_{L1} with the curves gives the values of I_a flowing for different values of V_g , e.g. when $V_g = -4$ volts, $I_a = I_{a4}$ when $V_g = -3$ volts, $I_a = I_{a3}$, etc. Plotting these values of I_a against V_g gives the dynamic characteristic of the valve with load R_{L1} (fig. 68 (b)). The dynamic characteristic for anode load R_{L2} (fig. 68 (c)) can be derived from fig. 68 (a) in the same way.

146. Examination of the curves shows that the dynamic characteristics of a pentode differ very little from the static $I_a:V_g$ characteristics provided anode bottoming does not occur, the current taken being practically independent of the load. However, it can be seen from the dynamic characteristic for R_{L2} that anode bottoming causes the curve to bend over and become horizontal at X , since further increase of V_g has no effect on I_a .

147. In a triode, I_a increases with V_g as it is varied from negative to positive values, and so the dynamic characteristic takes the same form as that for a pentode with a load which is too small for anode bottoming to occur (fig. 68 b). However, the anode current flowing in a triode with anode load for some particular

value of V_g is less than the corresponding static current. Thus the dynamic characteristics have a smaller slope than the $I_a:V_g$ static characteristic.

Pentode valve as a constant current generator

148. It may be seen from the pentode $I_a:V_a$ curves of fig. 69 that if V_g is maintained at some voltage level, then the variation of V_a over the range V_{a1} to V_{a2} (i.e. above the knee of the characteristic) causes very little variation in the anode current flowing so that the valve generates current of constant amplitude. If V_a and V_g both vary, then, provided that V_a does not fall below the knee of the characteristic, the changes in I_a are determined almost entirely by the variation of V_g .

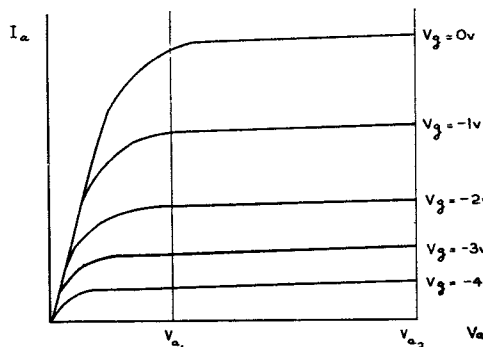


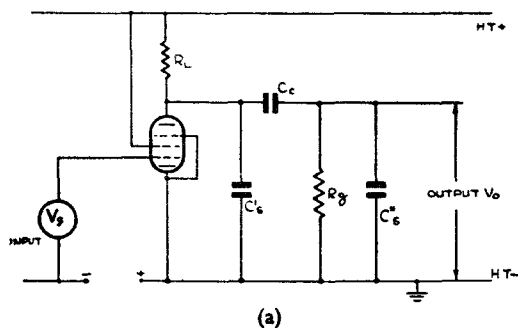
Fig. 69.—Pentode characteristic showing the region where V_a has little control

Resistance-capacitance (RC) coupled amplifier

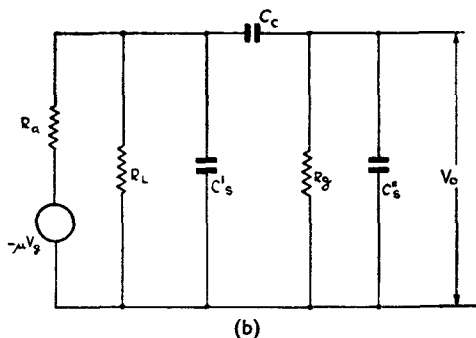
149. When RC coupling is employed, the output or anode voltage variation is passed on to the next amplifier stage by means of a blocking capacitor C_c and the grid leak resistance R_g of the next stage, fig. 70(a). Since it is required to develop practically all the output voltage across R_g , the capacitor C_c must be chosen so that its reactance is very small in comparison with R_g down to the lowest frequency which it is required to amplify. Thus, as far as AC variations are concerned, R_g is effectively in parallel with the anode load R_L . In order to prevent the effective anode load being reduced appreciably with consequent reduction of amplification by the stage, it is necessary to choose a grid leak such that $R_g \gg R_L$. But the maximum value of R_g is limited by the possibility of the flow of reverse grid current in the next stage

due to either (a) grid emission or (b) positive ion current due to traces of gas in the valve. The maximum safe value of grid leak resistance is of the order of 2M for small valves, and 100K for larger power valves.

150. In radar equipments, the main use of RC coupled amplifiers is to amplify DC pulses, e.g. the detector output in the superheterodyne receiver, the RF and IF amplifiers having tuned circuits as loads and transformer coupling between stages. Amplifiers with resistive loads and RC coupled to other stages are used also in the timing sections of equipments to handle square and pulse waveforms in addition to sawtooth time base waveforms.



(a)



(b)

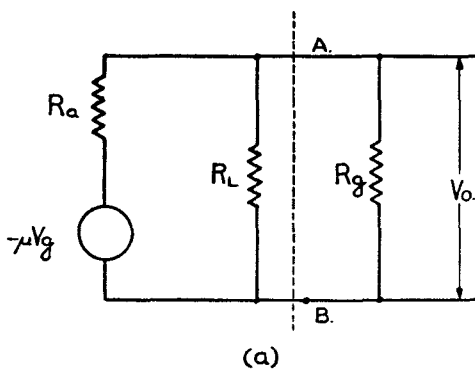
Fig. 70.—RC coupled amplifier

151. In each case it is possible to design the amplifier by two methods:—

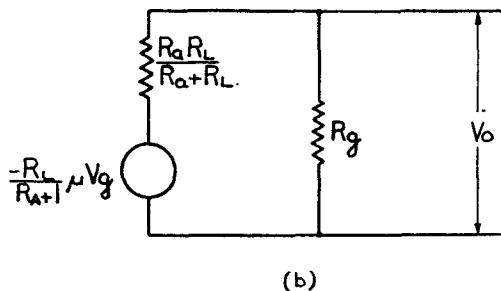
(i) From calculations determining the bandwidth required to pass with equal amplification and phase change all the sinusoidal components of the waveform.

(ii) From direct considerations of the circuit components which affect the rate of rise or fall of the input voltage, and those which cause variation in voltage when it is required to keep the output at a fixed voltage level for a certain interval of time.

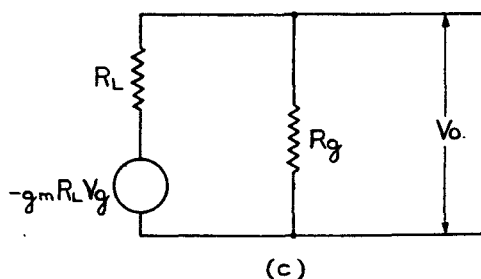
Whichever method is adopted, it is necessary to consider the stray capacitance in the circuit



(a)



(b)



(c)

Fig. 71.—Equivalent diagrams of RC amplifier for middle band of frequencies

C'_s between the anode and earth due to the anode-cathode capacitance of the valve and the wiring capacitance and C''_s across R_g due to the input capacitance of the next stage fig. 70 (a). C'_s should not exceed 10–15 pF, but the value of C''_s will vary between fairly wide limits, depending on the nature of the next stage. If the output is fed directly to the deflector plates of an oscilloscope, C'_s may be of the order of 100 pF, while if R_g is the grid leak of a pentode amplifier C''_s may be only 5 pF. If the valve is replaced by the equivalent voltage generator circuit (generator $-\mu V_g$ in series with the internal resistance R_a of the valve), then fig. 70 (b) gives the equivalent AC diagram of the amplifier.

Frequency response of the RC coupled amplifier

152. The output voltage tends to fall at low frequencies where the reactance of the coupling capacitor C_c becomes very large, and at high frequencies where the shunt reactance of the stray capacitance becomes finite.

153. For a well-designed amplifier there is a middle band of frequencies for which the reactance of C_c is so small that C_c can be regarded as a short circuit, while the reactances of C'_s and C''_s are so high that they have no appreciable shunting effect on R_L and R_g . The equivalent circuit then reduces to that shown in fig. 71 (a). This diagram can be further simplified to that of fig. 71 (b) by use of the following theorem.

Thevenin's theorem

154. Any linear network containing one or more sources of voltage and having two terminals, can be replaced by a generator of voltage E in series with an impedance Z where

E = Voltage that appears across the terminals when no load impedance is connected across them.

Z = Impedance between the terminals when all sources of voltage in the network are short-circuited.

In fig. 71 (a), the part of the network to the left of the dotted line is simplified by means of this theorem, R_g being considered as the load impedance across the terminals AB. Thus

$$Z = \frac{R_a R_L}{R_a + R_L} \quad (\text{i.e. } R_a \text{ and } R_L \text{ in parallel}).$$

$$E = \frac{R_L}{R_a + R_L} \cdot (-\mu V_g).$$

If a pentode valve for which $R_a \gg R_L$ is used, then Z will be very nearly equal to R_L , and

$$E \simeq \frac{R_L}{R_a} (-\mu V_g)$$

$$= -g_m R_L V_g.$$

giving the equivalent diagram of fig. 71 (c).

Since R_g is chosen to be very much greater than R_L , V_o will be nearly equal to the voltage $-g_m R_L V_g$ of the equivalent generator. The output voltage will be in phase with that of the equivalent voltage generator, because the circuit for this band of frequencies has only resistive elements.

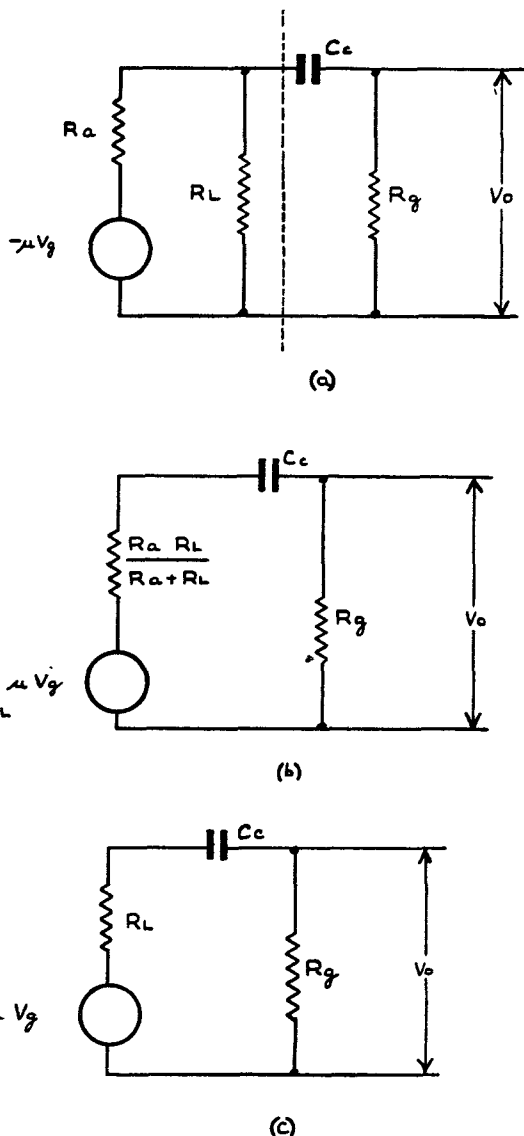
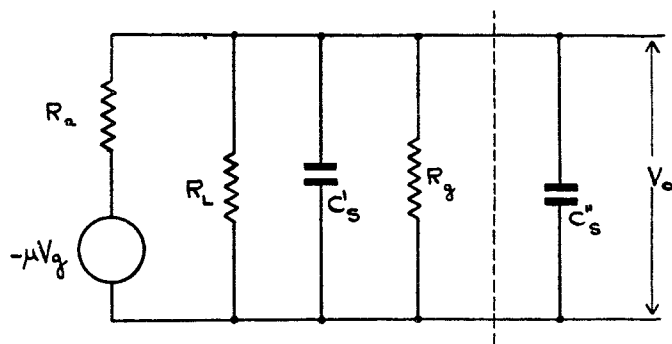


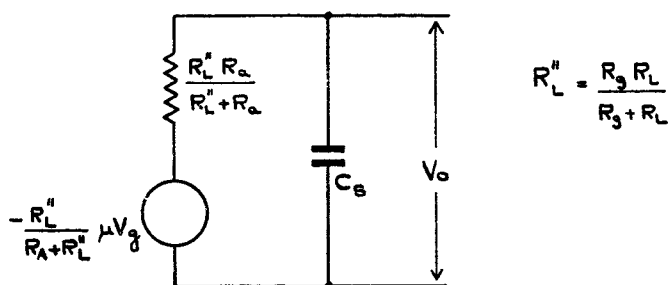
Fig. 72.—Equivalent diagrams of RC amplifier for low frequencies

Low frequencies

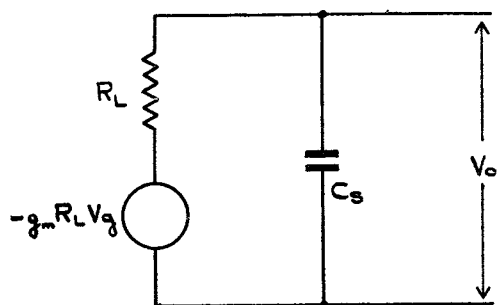
155. At frequencies below this middle band the reactance X_{C_c} of C_c increases to an appreciable value, while the reactances $X_{C'_s}$, $X_{C''_s}$ are too high to have any shunting effect. Thus the generalised equivalent diagram of fig. 70 (b) reduces to that of fig. 72 (a), further simplification by Thevenin's Theorem giving the network of fig. 72 (b). Assuming that a high R_a pentode valve is used, the equivalent circuit of fig. 72 (c) is obtained.



(a)



(b)



(c)

Fig. 73.—Equivalent diagram of RC amplifier for high frequencies

When $X_{Cc} \approx 0$, $V_o \approx \frac{R_g}{R_L + R_g} \cdot (-g_m R_L V_g)$
 $\approx -g_m R_L V_g$ if $R_g \gg R_L$.

i.e. the value of V_o obtained for the middle band of frequencies.

At lower frequencies

$$V_o = \frac{R_g}{\sqrt{(R_L + R_g)^2 + X_{Cc}^2}} \cdot (-g_m R_L V_g)$$

$$\frac{R_g}{\sqrt{R_g^2 + X_{Cc}^2}} \cdot (-g_m R_L V_g)$$

$$= \frac{1}{\sqrt{1 + \left(\frac{X_{Cc}}{R_g}\right)^2}} \cdot (-g_m R_L V_g).$$

At some frequency f_0 where $X_{Cc} = R_g$.

$$V_o = \frac{1}{\sqrt{2}} \cdot (-g_m R_L V_g).$$

$$\approx 0.7 \times (\text{output voltage for medium frequencies}).$$

(i.e., 3 db below maximum).

Below f_0 , X_{Cc} increases, and consequently V_o decreases further.

High frequencies

156. At high frequencies, X_{Cc} is small enough to be regarded as a short, but the shunting effect of the finite reactances $X_{C's}$ and $X_{C''s}$ now has to be considered. The equivalent circuit for high frequencies takes the form shown in fig. 73 (a) simplified by Thevenin's theorem in fig. 73 (b), and modified for the case of a high R_a pentode amplifier in fig. 73 (c). In this last diagram, it has been assumed that $R_g \gg R_L$ so that R_L and R_g in parallel (R''_L) are effectively equal to R_L .

The diagram indicates that the maximum value of V_o is $-g_m R_L V_g$ where the reactance of C_s ($= C's + C''s$) is very much higher than R_L .

Since $V_o = \frac{X_{Cs}}{\sqrt{R_L^2 + X_{Cs}^2}} \cdot (-g_m R_L V_g)$

$$= \frac{1}{\sqrt{1 + \left(\frac{R_L}{X_{Cs}}\right)^2}} \cdot (-g_m R_L V_g)$$

V_o falls to $\frac{1}{\sqrt{2}} \times (\text{output voltage at middle frequencies}).$
 (i.e. 3 db below maximum)

at the frequency f_1 where $X_{Cs} = R_L$. At higher frequencies V_o falls off until X_{Cs} is negligible in comparison with R_L and V_o is approximately zero.

157. In order that the amplifier may pass a non-sinusoidal waveform without distortion, it is necessary that the frequency spread of the waveform shall lie within the middle band of frequencies for which the gain of the amplifier is constant and the phase shift is zero. To make the middle band of frequencies extend over a large range,

(i) C_c must be large so that the frequency at which $X_{Cc} = R_g$ is low, and

(ii) R_L must be low, and C_s kept as small as possible so that the frequency at which $X_{Cs} = R_L$ is high.

Thus to get uniform amplification up to very high frequencies the anode load must be very low, with consequent low amplification in the stage.

158. *Example.*—Determine the range of frequencies for which the response does not fall below 3 db of the maximum for an amplifier with $R_L = 20K$, $R_g = 1M$, $C_c = 0.01 \mu F$, $C_s = 10 \text{ pF}$. $C''_s = 40 \text{ pF}$.

The lower frequency (f_0) at which the response is 3 db below maximum is given by

$$\frac{1}{2\pi f_0 C_c} \approx R_g.$$

$$\frac{1}{2 \times 3.142 \times 0.01 \times 10^{-6} \times f_0} \approx 10^6$$

$$f_0 \approx \frac{100}{6.286} \text{ c/s.}$$

$$\approx 16 \text{ c/s.}$$

Upper frequency (f_1) at which the response is 3db below maximum is given by

$$\frac{1}{2\pi f_1 C_s} \approx R_L.$$

$$\frac{1}{6.286 \times 50 \times 10^{-12} f_1} \approx 20 \times 10^3.$$

$$f_1 \approx \frac{10^6}{6.286}$$

$$\approx 160 \text{ Kc/s.}$$

The response curve for this amplifier is drawn in fig. 74.

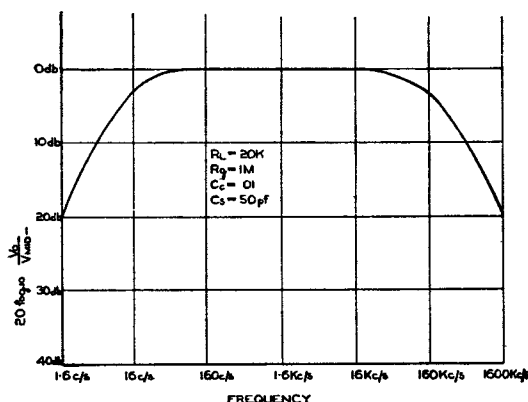


Fig. 74.—Frequency response curve for RC amplifier

Response to square waveform

159. Consider the application of an ideal square waveform to the grid of the amplifier shown in fig. 70 (a).

Edges of output waveform

160. When the grid voltage is raised instantaneously, the anode current increases, but the anode voltage falls only as rapidly as C'_s discharges, and the output voltage drop is also delayed by the discharge of C''_s . Similarly, when the input takes the grid more negative, the anode and output voltage changes are delayed by the charge of C'_s and C''_s . Since $C_c \gg C'_s + C_s$, it will charge or discharge by negligible amounts during the period required for the complete charge or discharge of the stray capacitance. Thus, when considering rapidly-changing input voltages, the effect of the coupling capacitor C_c on the output voltage can be neglected, and the equivalent amplifier circuit is derived by omitting C_c from fig. 70 (b). The equivalent circuits of fig. 73 are then obtained—the circuits already derived when considering a high frequency sine waveform input to the amplifier.

161. From fig. 73 (c) it can be seen that the output voltage cannot change instantaneously from one voltage level to another, but the change occurs in the time taken by the capacitor C_s to charge completely to the voltage of the equivalent generator, i.e. a time approximately equal to $5C_sR_L$. Thus, to make the edges as steep as possible, R_L must be low. It has already been seen that low R_L is the condition for good high frequency response. Fig. 75 (a) shows the delay in rise and fall of the edges of a square waveform due to the charge and discharge of the stray capacitance.

Flat portions of the output waveform

162. When the input takes the grid more positive, anode current increases and the voltage developed across R_L increases as rapidly as the strays discharge, i.e. V_a and the left-hand plate of C_c fall in voltage. Since C_c discharges by a negligible amount during this time, the voltage of its right-hand plate (V_o) falls by the same amount. During the period when the input voltage is held steady at its most positive value, C_c discharges, the anode voltage remaining approximately constant and the output voltage becoming less negative as it discharges. Since C''_s is very small, it discharges sufficiently rapidly to follow this change in output voltage. Similarly, when the grid is held at its most negative value, C_c charges up due to the rise in anode voltage, and as it does so the output voltage falls. Again, the strays follow the slow voltage change and so have no effect on the flat portions of the waveform. Thus the equivalent circuit to be considered when a steady voltage is applied to the grid is obtained by omitting C'_s and C''_s from fig. 70 (b). This gives the equivalent circuits already derived for a low-frequency sine waveform input (fig. 72). From fig. 72 (c), it can be seen that V_o is a maximum

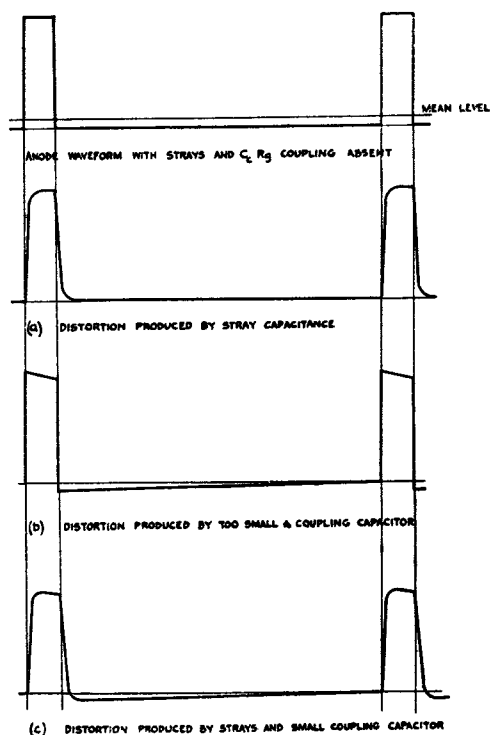


Fig. 75.—Distortion of pulse waveform by RC coupled amplifier

before C_c has begun to charge through R_L and R_g to the voltage of the equivalent generator. As C charges, V_o falls. In order to make the drop as small as possible during the period T for which the grid voltage is kept steady, the time constant $C_c (R_L + R_g) \approx C_c R_g$ must be very much larger than T . Fig. 75 (b) shows the distortion of the flat portions of the pulse due to the charge and discharge of the coupling capacitor. Fig. 75 (c) shows the distortion due to both strays and coupling capacitor.

163. *Example.*—Choose values of R_L , C_c and R_g for an RC coupled amplifier which is required to handle 100 μ sec negative pulses of p.r.f. 400 c/s, the total stray capacitance in the circuit being 50 pF. A delay of 1/20th of the pulse width in the rise and fall of the pulse, and 1% drop in amplitude can be tolerated. Since the output pulse must rise to its maximum value in $\frac{1}{20} \times$ pulse width, then the largest value of R_L which can be used is given by

$$5C_s R_L = \frac{1}{20} \times 100 \mu\text{sec.}$$

$$5 \times 50 \times 10^{-12} R_L = \frac{1}{20} \times 100 \times 10^{-6}.$$

$$\begin{aligned} \text{i.e. } R_L &= \frac{10^{-4}}{20 \times 5 \times 50 \times 10^{-12}} \text{ ohms.} \\ &= \frac{10^5}{5} \text{ ohms.} \\ &= 20\text{K.} \end{aligned}$$

R_g must be very much larger than R_L . Let $R_g = 1\text{M}$.

Now C_c charges to the mean of the applied waveform.

Since the pulse repetition period = $\frac{1}{400}$ secs.
= 2500 μ secs.

$$\frac{\text{Duration of pulse}}{\text{Pulse repetition period}} = \frac{1}{25}$$

Therefore the voltage to which C_c charges is very little different from the lower voltage extreme (fig. 75) of the anode waveform. Thus when the pulse is applied to the grid, the effective charging voltage applied to C_c is very nearly equal to the amplitude of the pulse. Since only 1% drop in amplitude of the output

pulse is permissible, C_c must charge through no more than 1% of the effective charging voltage during the pulse. A capacitor charges through 1% of the applied voltage in a time

$$\approx \frac{1}{100} C_c R_g, \text{ so that the smallest value of } C_c$$

is given by

$$\therefore \frac{1}{100} C_c R_g = 100 \mu\text{sec.}$$

$$\frac{1}{100} C_c \times 10^6 = 100 \times 10^{-6}.$$

$$\begin{aligned} C_c &= \frac{100 \times 100 \times 10^{-6}}{10^6} \text{ farads} \\ &= 10^4 \times 10^{-6} \mu\text{F} \\ &= 10^{-2} \mu\text{F} \\ &= .01 \mu\text{F.} \end{aligned}$$

If a smaller value of R_g had been chosen, a larger C_c would be required. Conversely, a smaller value of C_c could be used, if a larger R_g had been chosen.

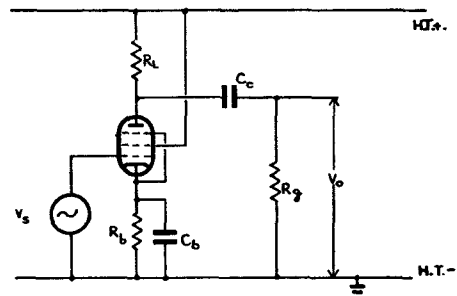


Fig. 76.—RC amplifier with self-bias arrangement

Biassing arrangements

164. If a self-biasing arrangement is used as shown in fig. 76, then the reactance of the bypass capacitor C_b must be low compared with the bias resistor R_b down to the lowest frequency which has to be amplified. Otherwise, the bias combination introduces negative feedback at the lower frequencies, i.e. the cathode voltage follows the grid, reducing the effective input between grid and cathode and so reduces the amplification for these frequencies. A non-sinusoidal waveform will be distorted if varying amounts of negative feedback occur for its component frequencies

165. Suppose the capacitor C_b is omitted, (fig. 77). Then negative feedback will occur

for all frequencies. When a signal V_s is applied to the grid, let voltage V_K appear across the cathode load due to the resulting variation in anode current, the grid cathode voltage being V_{gk}

Then

$$V_s = V_K + V_{gk}.$$

If a pentode is used the current change produced by a voltage change V_{gk} is $g_m V_{gk}$.

$$\therefore V_K = g_m R_K V_{gk}$$

$$\therefore V_s = V_{gk} (1 + g_m R_K).$$

Thus the effective signal $V_{gk} = \frac{V_s}{1 + g_m R_K}$

and hence the gain is reduced in the ratio $\frac{1}{1 + g_m R_K}$

e.g. if $R_K = 1K$ and $g_m = 5\text{ma/v}$

then the gain is reduced to $\frac{1}{1 + 5} = \frac{1}{6}$ th

of its value with no feedback.

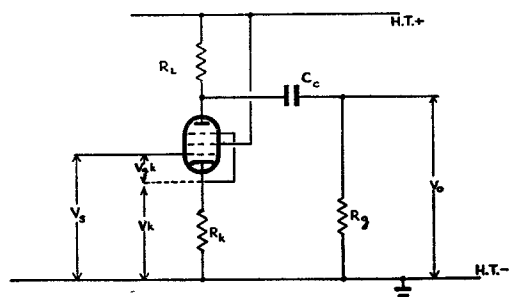


Fig. 77.—RC amplifier with unbypassed bias resistor

Inductance compensation

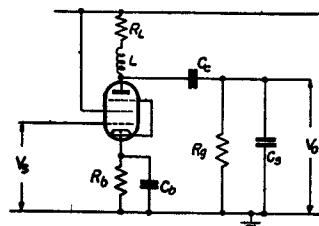
166. The response curve of the RC coupled amplifier can be maintained flat to a higher frequency by placing a small inductance L in series with the anode load (fig. 78 (a)). The cut-off at high frequencies is also made much sharper, i.e. the amplification falls off more rapidly beyond the flat region.

167. At low and medium frequencies the reactance $2\pi f L_C$ is small enough to be negligible, so that the response curve is the same as for the plain RC coupled amplifier. At high frequencies L tends to resonate with the stray capacities and raise the amplification. Response curves are drawn in fig. 78 (b) for several values of L , L_2 being the optimum value. Analysis shows that for pentode amplifiers, in order to obtain a flat response up to a particular frequency, the load R_L should equal the reac-

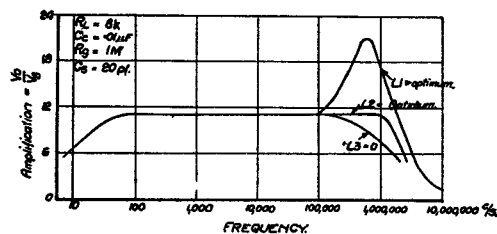
tance of the shunting capacities at that frequency, and the reactance of the inductance at the same frequency should equal half the load resistance, i.e.

$$\frac{1}{2\pi f_1 C_s} = R_L \quad \text{where } f_1 \text{ is the frequency up to which a flat response is required.}$$

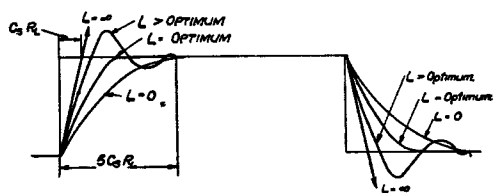
$$2\pi f_1 L = \frac{1}{2} R.$$



(a)



(b)



(c)

Fig. 78.—Inductance compensation

168. If the amplifier is required to handle pulse or square waveforms, the edges of the output waveforms will be made steeper by use of inductance compensation. It has been seen that for the RC coupled amplifier the time of rise and fall of the edges is approximately equal to $5C_s R_L$. This time of charge and discharge of the strays could be reduced to $C_s R_L$ provided some means of maintaining the initial rate of charge or discharge of C_s can be employed. Suppose an infinite inductance, i.e. one which allows no change in the current flowing through it were used to compensate the circuit. Then changes in anode current in the valve due to input voltage variations would alter only the current

I_c flowing in or out of the capacitor, the current I_L flowing through the inductance being unchanged. Thus C_s would charge to the maximum value of the pulse in time $C_s R_L$, but would continue to charge at this rate until the input voltage was switched, similarly for the discharge of C_s . If a practical inductance is used, the charging current will alter as the current I_L changes, but C_s will charge to the maximum value in some time between $C_s R_L$ and $5C_s R_L$. If the inductance is large and the current grows slowly in it there will be a tendency to overswing and produce ringing. The optimum value of L is that for which the ring is just damped out (fig. 78 (c)).

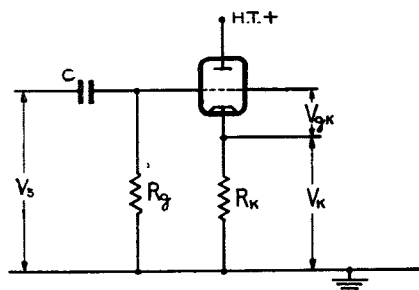


Fig. 79.—Basic cathode follower circuit

The cathode follower

169. The cathode follower is generally used as an impedance matching device since it has a high input and a low output impedance. The basic circuit of the cathode follower is shown in fig. 79. The anode is connected directly to the HT positive line (or to a decoupled anode resistor) while the resistor R_K in the cathode circuit is unbypassed. The output voltage is developed across R_K . If a positive signal is applied to the grid, the current in the valve increases and the voltage V_K across the cathode resistor increases. Similarly, if a negative signal is applied to the grid, the valve current decreases and V_K falls. Thus the cathode voltage "follows" the grid, and tends to decrease the voltage difference between the grid and cathode produced by the input signal.

Amplification

170. Fig. 80 gives the equivalent circuit of the cathode follower, the valve being replaced by its equivalent generator of voltage $-V_{gk}$ and internal resistance R_a . A positive signal V_s produces a positive voltage V_K across R_K . The voltage V_K opposes the input voltage,

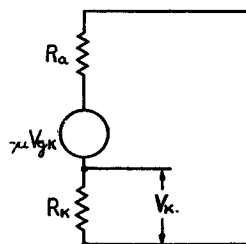


Fig. 80.—Equivalent circuit of cathode follower (1)

so that the actual voltage difference V_{gk} produced between the grid and cathode is equal to the difference between the input and output voltages.

$$\text{i.e. } V_{gk} = V_s - V_K.$$

From the equivalent diagram of fig. 80, we see that

$$V_K = \frac{R_K}{R_a + R_K} \cdot \mu V_{gk}$$

Substituting for V_K in the above expression

$$V_{gk} = V_s - \frac{R_K}{R_a + R_K} \cdot \mu V_{gk}$$

$$\frac{V_{gk}(R_a + R_K + \mu R_K)}{R_a + R_K} = V_s.$$

$$V_{gk} = \frac{R_a + R_K}{R_a + (\mu + 1) R_K} \cdot V_s$$

$$\approx \frac{R_a + R_K}{R_a + \mu R_K} \cdot V_s$$

where $\mu \gg 1$

$$V_K = \frac{\mu R_K}{R_a + R_K} \cdot V_{gk}$$

$$= \frac{\mu R_K}{R_a + R_K} \cdot \frac{R_a + R_K}{R_a + \mu R_K} V_s$$

$$= \frac{R_K}{R_K + \frac{1}{g_m}} \cdot V_s$$

dividing numerator and denominator by μ .

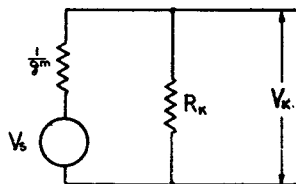


Fig. 81.—Alternative equivalent circuit for cathode follower

This expression for V_K indicates that the cathode follower is equivalent to a generator of voltage V_s with internal resistance $\frac{1}{g_m}$, the voltage being developed across the load R_K (fig. 81).

Provided $R_K \gg \frac{1}{g_m}$, V_K is approximately equal to V_s and the gain of the stage, i.e. $\frac{V_K}{V_s}$ differs little from unity. However, if the cathode resistor R_K is of the same order as $\frac{1}{g_m}$, then the output voltage is very much less than the input.

Handling capacity

171. The above theory and equivalent diagram are only true when the valve is working over the linear parts of its characteristics, and do not hold when the input signal cuts off the valve current or causes it to overload due to grid current flow. It may be seen that the maximum output voltage obtainable without overloading the valve, is larger the larger the cathode load.

$$\begin{aligned} \text{Now } V_{gk} &= V_s - V_K \\ &= V_s - \frac{R_K}{R_K + \frac{1}{g_m}} \cdot V_s \\ &= \left(\frac{\frac{1}{g_m}}{R_K + \frac{1}{g_m}} \right) V_s \\ &= \frac{1}{1 + g_m R_K} V_s. \end{aligned}$$

Then as R_K is increased, the fraction of the signal voltage which appears between the grid and cathode is reduced, which means that the handling capacity is increased. This may be illustrated by an example.

172. Assume that the valve has a g_m of 5 mA/volt and a grid base of 4 volts. Therefore, neglecting the curvature of the valve characteristics, the permissible variation of V_{gk} for undistorted output is 4 volts, i.e. from 0 (where grid current starts to flow) to -4 volts (where the valve current cuts off).

$$V_s = V_{gk} (1 + g_m R_K)$$

If an undistorted output is required, then for $R_K = 10K$

$$\begin{aligned} (V_s)_{\max} &= 4 (1 + 50) \text{ volts} \\ &= 204 \text{ volts} \end{aligned}$$

$$\begin{aligned} \text{For } R_K = 1K \quad (V_s)_{\max} &= 4 (1 + 5) \text{ volts} \\ &= 24 \text{ volts.} \end{aligned}$$

This indicates that by using a large cathode load, a large signal voltage can be handled, i.e. a large undistorted output is obtainable.

Biassing arrangements

173. In the basic circuit of fig. 79, self bias is provided by means of the load resistor R_K , the grid leak resistor R_g being returned to earth. In general, the negative bias produced in this way is too large and two arrangements which can be used to give reduced bias are shown in fig. 82, in each case R_g being returned to a voltage more positive than earth.

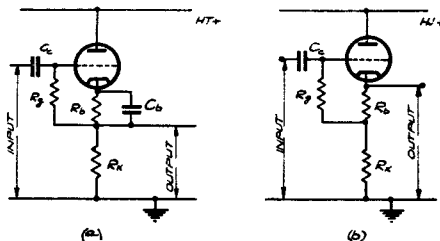


Fig. 82.—Biassing arrangements for the cathode followers

Output impedance

174. The output impedance of a cathode follower, i.e. the impedance which it presents to the following stage, is extremely low provided the valve is conducting. The equivalent diagram of fig. 81 can be transformed by Thevenin's theorem to a generator of voltage

$$\frac{R_K}{R_K + \frac{1}{g_m}} \cdot V_s \text{ in series with the resistance of } R_K \text{ and } \frac{1}{g_m} \text{ in parallel (fig. 83), i.e. the output impedance is } R_K \text{ in parallel with } \frac{1}{g_m}.$$

If $g_m = 5 \text{ mA/volt}$, $\frac{1}{g_m} = 200 \text{ ohms}$, and the output resistance is 200 ohms in parallel with R_K ; this can never exceed 200 ohms.

175. Its low output impedance makes the cathode follower useful for

(i) supplying a low impedance load (e.g. a length of matched cable)

(ii) feeding into a circuit with shunt capacitance (e.g. a short length of unmatched cable).

176. Suppose the stage fed by the cathode follower throws a capacitance C_K across the cathode load R_K (fig. 84). The output voltage can alter only as rapidly as C_K charges or

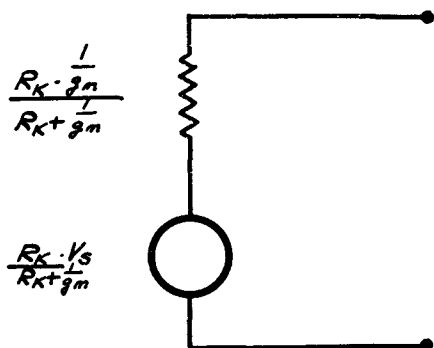


Fig. 83.—Equivalent circuit of cathode follower (2)

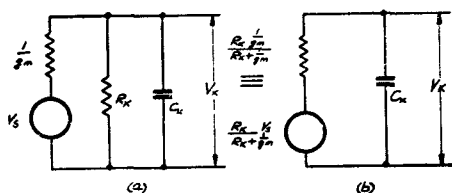


Fig. 84.—Equivalent circuit of cathode follower with capacitive load

discharges. From the equivalent generator diagram of fig. 84 (b) it may be seen that C_K charges and discharges through R_K and $\frac{1}{g_m}$ in parallel, to the voltage $\frac{R_K}{R_K + \frac{1}{g_m}} \cdot V_g$

of the equivalent generator. Since the effective resistance is very small, the time constant of charge and discharge of C_K is small, so pulse and square waveforms will be handled with little distortion.

From the point of view of frequency response, the shunting effect of C_s at high frequencies will be small since $\frac{1}{g_m}$ is very low, so the bandwidth will be large.

177. This is only true provided the valve is working over the linear part of its characteristics. Suppose a large amplitude positive pulse is applied to the grid of the cathode follower (fig. 85), the bias being such that when no input signal is applied, the cathode voltage V_K is +50 volts and the grid voltage V_g is +47 volts, i.e. V_{gk} is -3 volts. When the positive pulse is applied, V_g is raised from 47 volts to 97 volts and the valve current

grows rapidly. However, V_K does not rise instantaneously, since C_s must first be charged.

C_s charges rapidly through R_K and $\frac{1}{g_m}$ in parallel (fig. 84), and V_K rises to its maximum value, 98 volts, say. V_{gk} is now only -1v, since a smaller negative grid-cathode voltage is required to produce the larger current now flowing. When the trailing edge of the input pulse causes V_g to fall by 50 volts, V_K cannot follow instantaneously due to the presence of C_s , and so the valve current is cut off. Now C_s begins to discharge through R_K , causing V_K to fall. If the interval between successive positive pulses is sufficiently long, V_K will fall within the grid base of the valve before the next rising edge of the input occurs. This causes the valve to conduct again, rapidly completing the discharge of C_s to its initial value, i.e. 50 volts.

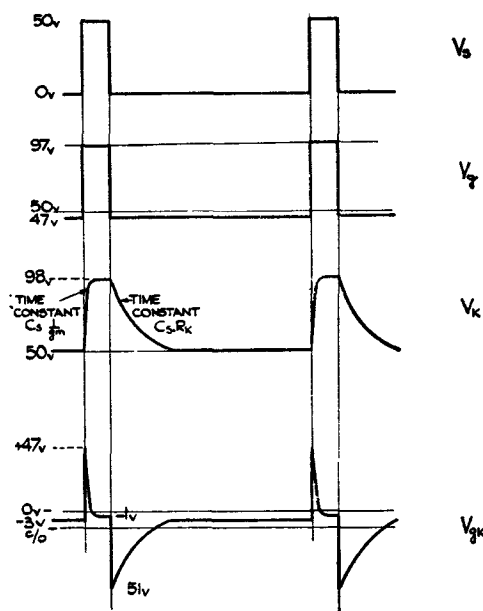
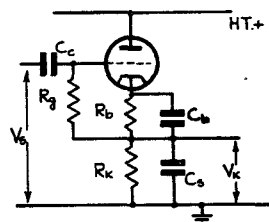


Fig. 85.—Distortion of positive pulses by shunt capacitance across R_K

178. Thus application of a 50-volt peak-to-peak positive pulse to the cathode follower produces an output pulse of slightly smaller amplitude, the positive edge of which rises on the time constant $C_s \times \frac{1}{g_m}$, while the negative edge falls on the time constant $C_s \times R_K$ (fig. 85). So, to make the trailing edge of the output pulse fall rapidly, R_K must be made low. For a low load, the amplification of the cathode follower is low, and so a compromise has to be made between amplification and preservation of pulse shape. The best results are obtained by using a low load for good pulse shape, with a heavy current valve to give the required amplitude of output waveform.

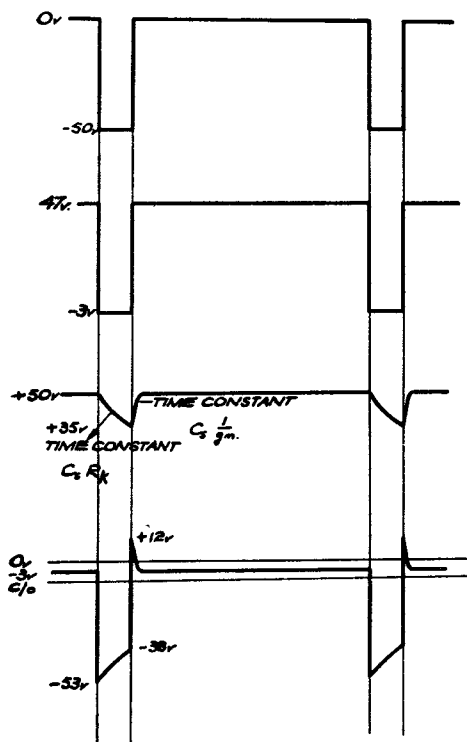
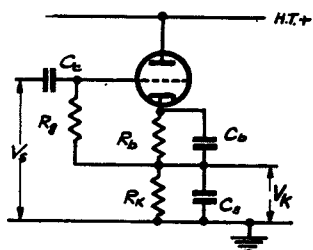


Fig. 86.—Distortion of negative pulses by shunt capacitance across R_K

179. It is more difficult to design a cathode follower to handle large amplitude negative pulses. The leading edge of the pulse now cuts off the valve current, and V_K falls on the time constant $C_s R_K$ as C_s discharges. If the pulse is narrow, and R_K large, V_K may not fall sufficiently to bring the valve into conduction before the trailing edge raises the grid voltage to its original value (fig. 86). Consequently, the pulse is badly distorted and the amplitude is reduced, being smaller the narrower the pulse.

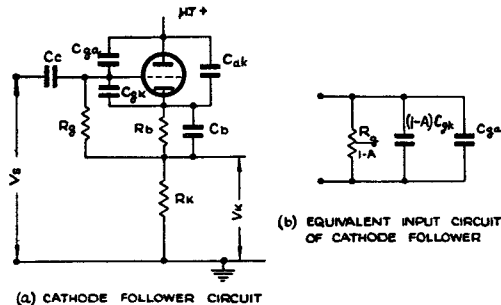


Fig. 87.—Cathode follower circuit

Comparison of the input impedance of the cathode follower and the RC coupled amplifier with resistive loads

Cathode follower

180. The input capacitance of a cathode follower is determined mainly by the inter-electrode capacitance C_{ga} , between the grid and anode, and to a lesser extent by the capacitance C_{gk} between the grid and cathode of the valve (fig. 87a). Since the anode is connected to the positive H.T. supply which is earthy as far as AC variations are concerned. C_{ga} appears to be directly across the input voltage source. C_{gk} also takes current from the source and its effect on the input impedance i.e. the impedance presented to the input source, will now be investigated.

Note.—The anode-cathode capacitance C_{ak} is directly across the load R_K and affects only the output impedance.

181. When a voltage V_s is applied between the grid and earth, a cathode voltage V_K is produced.

$$V_K = AV_s \quad \text{where } A \text{ is the amplification of the cathode follower.}$$

Thus the voltage change between the grid and cathode is

$$\begin{aligned} V_{gk} &= V_s - V_K \\ &= (1 - A) V_s. \end{aligned}$$

This voltage $(1 - A) V_s$ appears across C_{gk} and causes a current I_c to flow in it where

$$I_c = \frac{(1 - A) V_s}{\text{Reactance of } C_{gk}} \\ = (1 - A) V_s \cdot \omega C_{gk}.$$

This current is the same as that which would be taken from the source by capacitance $(1 - A) C_{gk}$ placed directly across it.

182. If the bias resistor R_b , is adequately bypassed, there will be no voltage variation across it due to the input voltage V_s , and so the total grid-cathode voltage change $(1 - A) V_s$ appears across R_g . The current taken from the source by R_g is thus

$$I_R = \frac{(1 - A) V_s}{R_g}$$

which is equal to the current which would flow from the source through resistance $\frac{R_g}{1 - A}$ placed directly across it.

Thus the valve input circuit appears to the preceding stage as resistance $\frac{R_g}{1 - A}$, and capacitance C_{ga} and $(1 - A) C_{gk}$ in parallel (fig. 87b). If R_K is very much higher than $\frac{1}{g_m}$ for the valve, A differs little from unity and $\frac{R_g}{1 - A}$ becomes infinitely high, while $(1 - A) C_{gk}$ tends to zero. So the input resistance of a cathode follower can be made extremely high, and the input capacitance little higher than the anode-grid capacitance of the valve.

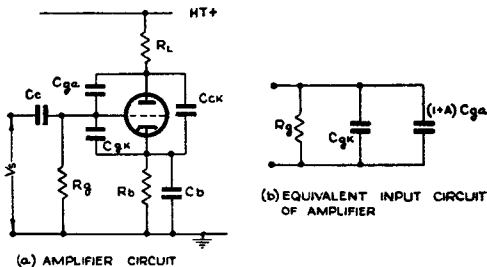


Fig. 88.—Amplifier circuit

183. For most pentodes C_{ga} , and consequently the input capacitance is considerably less than 1 pF; for a triode C_{ga} is of the order of 2 to 5 pF. Even the triode cathode follower circuit has an extremely low input capacitance, and it is preferred to the pentode circuit because of its greater simplicity.

Amplifier

184. When an input voltage V_s is applied to the grid of the amplifier (fig. 88a), the whole voltage variation appears across R_g , and across the capacitance C_{gk} provided that the bias resistor R_b is adequately by-passed. R_g and C_{gk} in parallel, however, do not form the total input impedance, since C_{ga} also takes current from the input voltage source. Since the anode load R_L is resistive, the anode voltage change is AV_s where A is the amplification of the stage. Thus the voltage developed across C_{ga} is $(1 + A) V_s$. The current flowing into C_{ga} from the source is

$$I_c = \frac{(1 + A) V_s}{\text{Reactance of } C_{ga}} \\ = (1 + A) V_s \cdot \omega C_{ga}.$$

which is equal to the current which would flow from the source into capacitance $(1 + A) C_{ga}$ placed directly across it. Thus, the input circuit of the amplifier appears to the preceding stage as a resistance R_g in parallel with capacitance C_{gk} and $(1 + A) C_{ga}$ (fig. 88b). Since A is always greater than unity, the input capacitance of the amplifier is greater than that of the cathode follower.

185. *Example.*—If an amplifier stage using a triode with $C_{gk} = 5$ pF and $C_{ga} = 4$ pF has an amplification of 50, then the input capacitance of the stage is

$$C_{gk} + AC_{ga} = 5 + 200 \\ = 205 \text{ pF.}$$

Suppose a pentode with $C_{gk} = 11$ pF, and $C_{ga} = 0.006$ pF is used, the amplification of the stage now being 100.

Then, input capacitances

$$= C_{gk} + AC_{ga} \\ = 11 + 0.6 \\ = 11.6 \text{ pF.}$$

Even though a stage with higher amplification and high grid-cathode capacitance has been chosen, the input capacitance is very much less than that of the triode amplifier considered. Hence it is usual to employ pentode valves in amplifier circuits where low input capacitance is required, e.g. the VF amplifier stages of the radar receiver.

RESISTANCE—CAPACITANCE OSCILLATORS

186. Low frequency oscillators are used as master oscillators in the control or timing sections of many radar equipments. Simple LC oscillators of the Hartley type are suitable for medium and high audio frequencies since the coils and capacitors required are small and can be constructed to have low losses. At the very low frequencies, such oscillators are impracticable, since it is difficult to reduce the losses associated with large inductances. Beat frequency oscillators will produce low frequency oscillations, but suffer from the following disadvantages:—

- (i) The frequency stability is poor, since a small change in frequency of one oscillator produces a large percentage change in the beat frequency.
- (ii) The oscillators have to be well shielded to prevent synchronisation at low frequencies.
- (iii) Calibration is not constant and has to be checked frequently.

These difficulties can be overcome by using a resistance capacitance (RC) oscillator. One type of RC oscillator which is commonly used, the Phase Shift Oscillator (PSO) will be described.

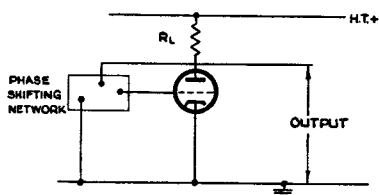


Fig. 89.—Base circuit of the phase shift oscillator (PSO).

Phase shift oscillator

187. The phase shift oscillator consists of a single amplifier valve with a phase-shifting network connected between its anode and grid, which introduces a phase-change dependent on the frequency passed through it (fig. 89). In order to build up or maintain oscillations in the circuit the voltage fed back from the anode to the grid must be phase-shifted by 180 deg. so as to be in phase with the original grid voltage, since the valve introduces 180 deg. phase shift between grid and anode. It is also essential for the production of self-sustained oscillations that the overall amplification of the stage shall be greater than or equal to unity:—

i.e. if A = voltage gain of the amplifier

β = fraction of the output voltage of the amplifier fed back to the input of the amplifier

then

$$A\beta > 1.$$

In the PSO the phase-shifting network takes the form of a number of RC networks in series. Application of an alternating voltage V_s to the single RC network shown in fig. 90 (a) causes a current to flow in the circuit of magnitude determined by the total impedance in the network. This current leads in phase on the applied voltage since the impedance is capacitive. The voltage V_R developed across R is in phase with the current flowing while the voltage V_C across C lags by 90 deg. on the current. Thus V_R leads the applied voltage by an angle θ , and V_C lags by an angle $(90-\theta)$, where $\tan \theta = \frac{\text{reactance of } C}{R} = \frac{1}{2\pi fCR}$

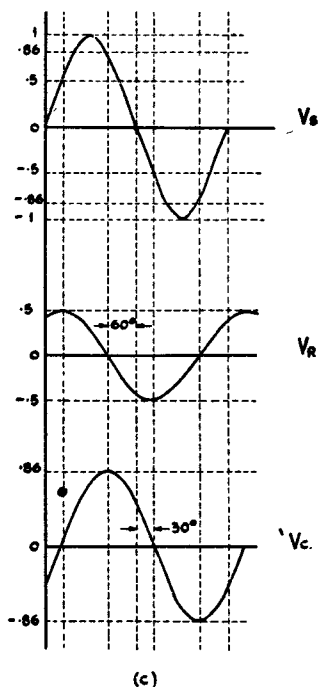
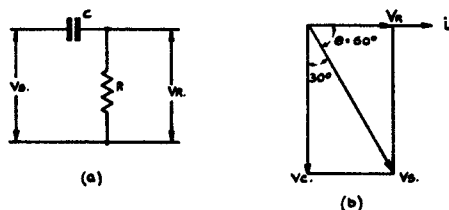


Fig. 90.—RC phase-shifting network

(fig. 90b). The vector diagram and voltage waveforms of fig. 90 (b) and (c) are drawn for values of C and R which, at the particular frequency of the input, are such as to give a voltage V_R 60 deg. ahead of the applied voltage. If R is decreased, the phase angle of V_R is increased, being 90 deg. when R is reduced to zero. This is useless since there is then no resistance across which to develop a useful voltage. Hence it is impossible to obtain 90 deg. phase shift from a single RC network of this type, and in order to obtain a phase shift of 180 deg. it is necessary to connect three or more RC sections in series.

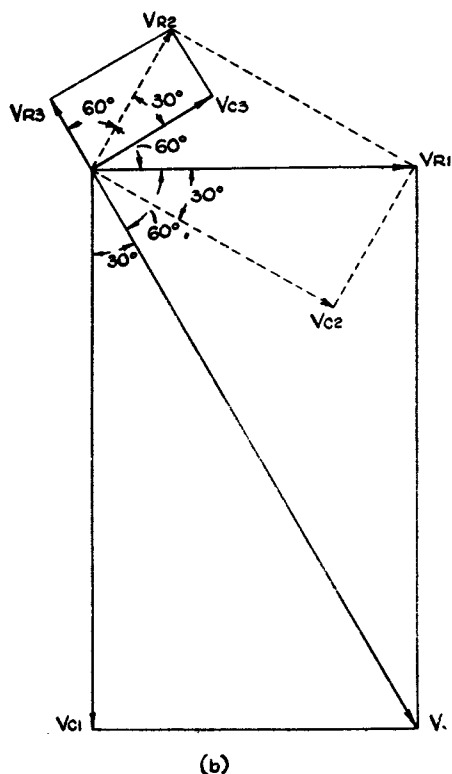
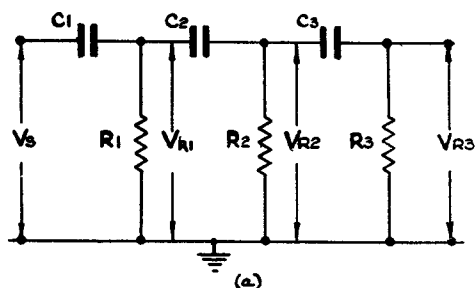


Fig. 91.—3-mesh RC phase-shifting network

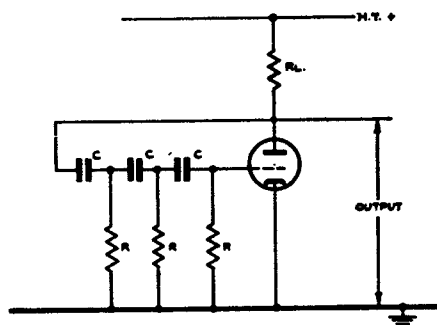
188. Consider the application of an alternating voltage V_s to a network consisting of three RC meshes connected in series as in fig. 91 (a). Due to the application of V_s to the first section, a voltage V_{R1} is developed across R_1 which leads in phase on V_s . V_{R1} is the effective voltage input for the second section and produces a voltage V_{R2} across R_2 which leads on V_{R1} . Similarly, application of V_{R2} to the third section develops an output voltage V_{R3} leading on V_{R2} . Consequently, V_{R3} leads V_s by a phase angle which is equal to the sum of the phase changes produced by each of the meshes of the network. For one particular frequency, the total phase change along the network will be equal to 180 deg.; for lower frequencies the phase change will exceed 180 deg.; for higher frequencies the phase change will be less than 180 deg. The output voltage differs not only in phase, but also in amplitude from the input, each section producing an output voltage smaller in amplitude than that applied to it. This can be seen from the vector diagram of fig. 91 (b), which is drawn for a three-mesh network, each mesh producing 60 deg. phase lead at the frequency, which gives 180 deg. total phase change.

189. The phase-shifting network of the PSO takes the form of three (or more) RC meshes, as in fig. 92, and so only a small fraction of the output voltage is fed back to the grid, but oscillations, once started, will be maintained, provided the amplification of the tube is sufficiently large. The oscillations are started by any circuit change such as switching on the HT supply or random valve noise, which produce slight voltage variations at the grid comprising all frequencies. This is amplified, phase-changed by 180 deg. from grid to anode, and fed back to the grid via the RC network. One frequency is phase-shifted 180 deg. by the network, returns to the grid in phase with the original variation and is re-amplified. This cumulative build-up is repeated until the valve cannot amplify further, and then the oscillations are maintained at constant amplitude.

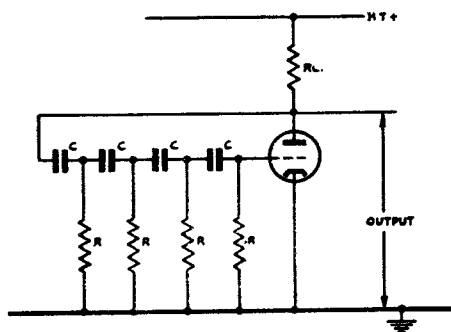
190. Harmonics of the fundamental frequency of the PSO are phase-shifted by only a small amount and suffer little attenuation in passing through the RC network; consequently, the feedback becomes negative. This prevents excessive amplification of harmonics and tends to produce a pure sine wave output, provided

the amplification of the valve is such that oscillations are just maintained.

191. The PSO is not generally required to produce a variable frequency output, but this can be achieved by making either the resistors or capacitors variable. The phase lead introduced by each RC section is determined by the ratio $\frac{\text{reactance of } C}{R} = \frac{1}{2\pi fCR}$. For oscillations to occur this ratio must have one particular value. Thus if it is required to increase the frequency of oscillation, either C or R must be decreased to keep the ratio constant. Similarly, to decrease the frequency, C or R must be increased. If only a limited frequency range is required, then only one resistor or capacitor need be made variable.



(a) 3 MESH PHASE SHIFT OSCILLATOR



(b) 4 MESH PHASE SHIFT OSCILLATOR

Fig. 92.—Phase shift oscillators with phase-advancing networks

192. It can be shown by analysis of the three-mesh PSO, that if the meshes are identical, the frequency of oscillation is $\frac{1}{2\pi\sqrt{\frac{10}{7}}RC}$, while the amplification required is 5.5.

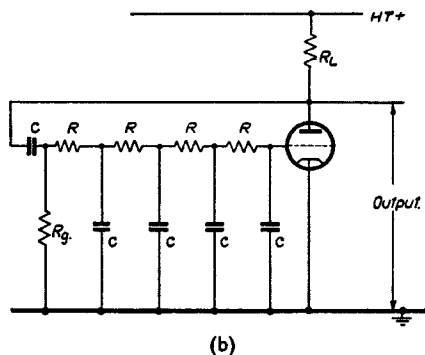
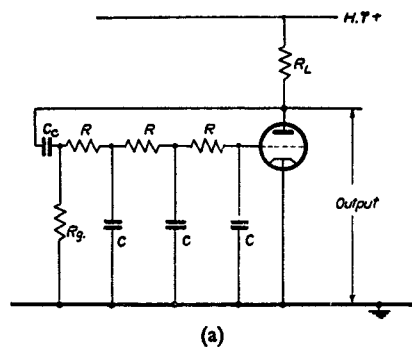


Fig. 93.—Phase shift oscillators with phase-retarding networks

The PSO circuits of fig. 93 have phase-shifting networks which produce a phase lag of 180 deg. Such a circuit with three identical meshes oscillates at a frequency $\frac{\sqrt{6}}{2\pi RC}$, the required amplification being 5; with four identical meshes the frequency is $\frac{\sqrt{10}}{2\pi RC}$ and the required amplification 18.4.

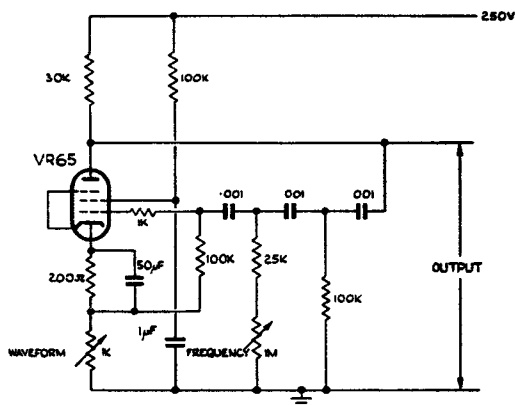


Fig. 94.—Phase shift oscillator circuit

193. Fig. 94 gives the circuit of a three-mesh PSO designed to work at approximately 400 c/s. The frequency can be varied above and below this value by means of the variable resistor R2. The unbypassed variable resistor R1 makes a waveform control. If the voltage fed back to the grid is so large that the output waveform is distorted, R1 is set at such a value that the voltage variation developed across it reduces the effective grid cathode voltage and causes the valve to work over the linear part of the characteristics, i.e. R1 introduces a variable negative feedback control. This negative feedback reduces the gain, but gives a distortionless sine waveform.

194. The advantages of the PSO as a low-frequency oscillator are:—

- (i) It gives a good waveform, particularly with negative feedback.
- (ii) It has good frequency stability.
- (iii) It is cheap and compact.
- (iv) The frequency is easily variable.
- (v) The upper frequency limit is of the order of 40 kc/s, while there is no lower frequency limit.

GENERATORS OF SQUARE AND PULSE WAVEFORMS

The squarer

Grid current effects

195. In normal radio practice it is customary to avoid grid current but in radar circuits grid current is often used to give the desired result. In a pentode valve the curve of grid current against grid volts depends to a certain extent on the voltage applied to the screen; further, the flow of current starts when the grid is somewhat negative with respect to the cathode and the curve is far from linear.

196. Since it would be extremely difficult to take all these effects into account, the following simplifying assumptions will be made:—

- (i) Grid current starts at zero grid volts.
- (ii) The grid current characteristic is linear and has a slope corresponding to a resistance of about 1K.

Between grid and cathode, we therefore have in effect a diode with a forward resistance of 1K and an infinite resistance in the reverse direction.

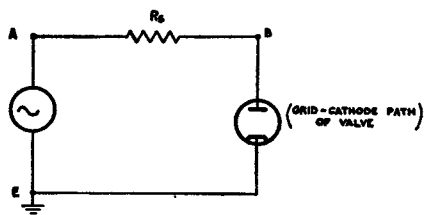


Fig. 95.—Circuit illustrating action of grid stopper

The grid stopper

197. Consider the circuit of fig. 95 when the point A is positive with respect to E. The diode is conducting with an effective resistance R_v . The voltage across the diode is therefore less than that across AE in the ratio $\frac{R_v}{R_s + R_v}$.

When A is negative with respect to E the diode is not conducting and the voltage at B is equal to that at A. If the generator delivers a sine wave output one cycle of the respective waveforms is shown in fig. 96. The positive

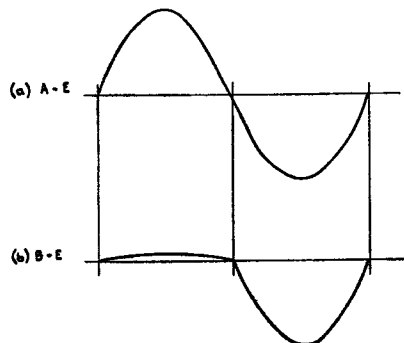


Fig. 96.—Voltage waveforms of fig. 95

half-cycle in fig. 96 (b) has been exaggerated; with $R_v = 1K$, $R_s = 1M$ its amplitude is only $\frac{1}{1000}$ th of the negative half-cycle.

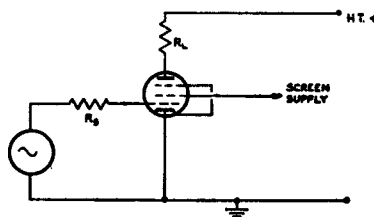


Fig. 97.—Simple squarer stage without bias

Squarer without bias

198. Fig. 97 shows the simplest form of

squarer stage. The anode load R_L is so chosen that the valve is bottomed at or below zero grid volts.

199. Fig. 98 shows the operation of the circuit and is almost self-explanatory. The dynamic characteristic bends over and becomes almost horizontal at X so that a further rise of grid voltage beyond this point has a negligible effect on the anode current. The positive half-cycle from the generator is very much reduced in amplitude by grid current in conjunction with R_s and consequently this half-cycle is represented by the portion AB of the grid voltage waveform.

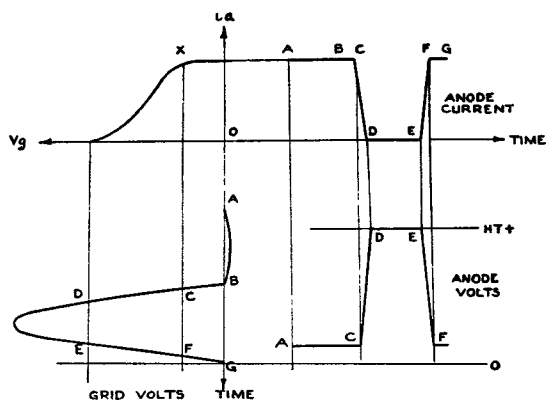


Fig. 98.—Waveforms of circuit of fig. 97

200. Throughout one cycle the voltages and currents are thus:—

	<i>Grid volts</i>	<i>Anode current</i>	<i>Anode volts</i>
AB	Positive swing very small due to grid current action.	Steady at bottoming value.	Steady at bottoming value (about 10 volts).
BC	Falling rapidly.	As above.	As above.
CD	Falling rapidly and approaching cut-off at D.	Falling rapidly along steep part of dynamic characteristic.	Rising rapidly to HT + as i_a falls.
DE	Below cut-off.	Zero.	HT +.
EF	Rising rapidly.	Rising rapidly on steep part of dynamic characteristic.	Falling rapidly as i_a increases.
FG	Rising rapidly.	Steady at bottoming value.	Steady at bottoming value.

201. It will be noted that the duration of the positive-going portion of the anode voltage waveform is slightly less than a half-cycle and that a perfect square wave is not produced at the anode since CD and EF are not vertical. This is because we have chosen an input which is not very much greater than the grid base. Clearly, if the input were, say, 20 times the grid base the grid voltage would sweep across the grid base at a very high speed and the steepness of the edges of the anode voltage waveform would be correspondingly improved.

202. If the anode load R_L were reduced sufficiently or if a triode valve were used, the dynamic characteristic would not bend over and become horizontal as at X in fig. 98. The portion AB of the grid voltage waveform would then appear in an amplified and inverted form on the anode voltage waveform (fig. 99).

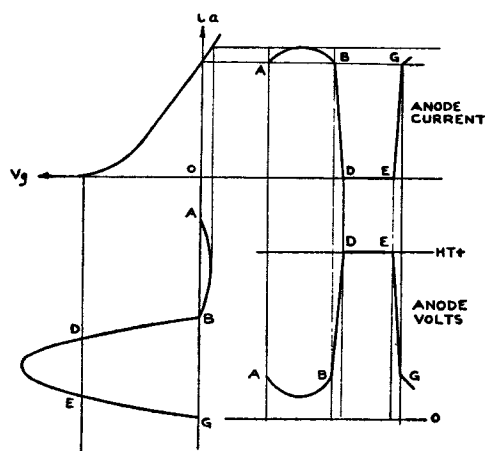


Fig. 99.—Modification of waveforms of fig. 98 when no bottoming action occurs

The design data for efficient squaring are thus:—

- (i) A short grid-base pentode to give steep edges to the output waveform.
- (ii) An input of large amplitude (of the order of 100 volts peak).

(iii) A high value grid stopper to remove the positive swing at the grid.

(iv) An anode load high enough to give effective bottoming.

Squarer with bias

203. Without bias we are limited to an output square wave with very nearly equal positive and negative-going portions. For the production of asymmetrical square waves bias is essential.

(a) Negative bias

The lettering of fig. 100 corresponds with that of fig. 98, and it is clear that the use of negative bias has reduced the duration of the negative-going portion of the waveform.

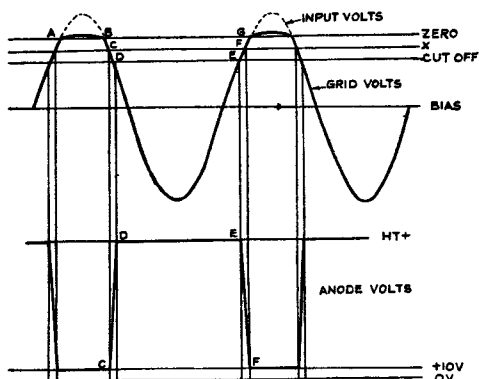


Fig. 100.—Waveforms of squarer with negative bias

(b) Positive bias

The corresponding case for +ve bias is shown in fig. 101. Here the positive-going portion of the output waveform is considerably shorter than the negative-going.

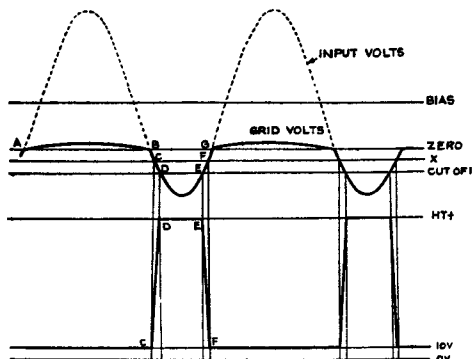


Fig. 101.—Waveforms of squarer with positive bias

204. Since the slope of a sine waveform decreases as the peak is approached, the rate at which the grid voltage moves over the

portion CD also decreases. When a large bias, either positive or negative, is used in an attempt to produce a very asymmetrical square waveform, some deterioration in the steepness of the edges of this waveform is therefore inevitable.

205. Fig. 102 shows a circuit enabling positive or negative bias to be selected at will. The point A is at about +160 volts and the point B, to which the cathode is connected, at about +80 volts. The setting of P therefore enables the bias to be varied between -80 volts and +80 volts.

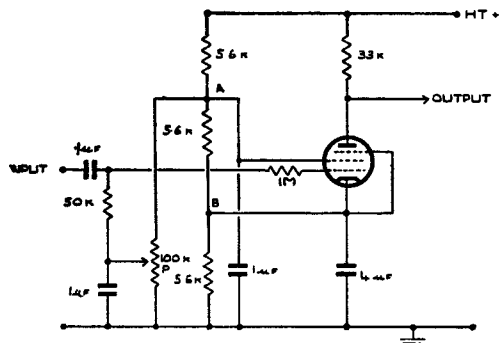


Fig. 102.—Circuit of squarer stage with arrangement for positive or negative bias

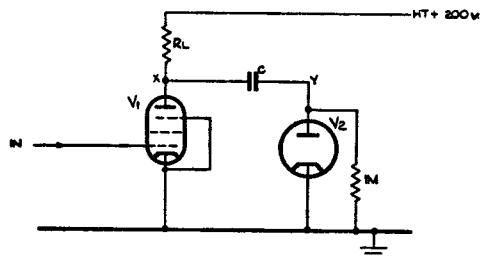


Fig. 103.—Subsidiary circuit illustrating charging of coupling capacitor C

The Multivibrator

206. Consider the circuit of fig. 103 and suppose that the grid is held at zero for a considerable period. If R_L is large enough for V_1 to be bottomed at zero grid volts the voltage at X will then be +10 volts and that at Y zero. The p.d. across the capacitor is thus 10 volts. Now suppose V_1 to be suddenly cut off by a large negative edge on the grid. C is now free to charge through R_L and the diode V_2 . The $1M$ in shunt with V_2 may be neglected since V_2 when conducting has a

resistance of about 1K. Since the capacitor voltage cannot change instantaneously from its original value of 10 volts, the remaining 190 volts must be shared between R_L and V_2 and the voltage distribution at the instant after V_1 is cut off is therefore as shown in fig. 104. As C charges the potential of X rises to $HT+$ and that of Y falls to zero, the time constant in each case being $C(R_L + R_v)$ or CR_L approx. since $R_v \ll R_L$.

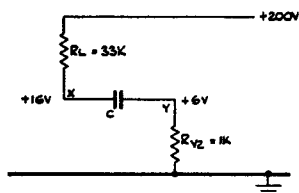


Fig. 104.—Effective circuit of fig. 103 when V_1 is suddenly cut off

207. The waveforms are shown in fig. 105. This subsidiary circuit occurs in the multi-vibrator and this preliminary study enables us to obtain a very much clearer picture of certain aspects of the operation of that circuit.

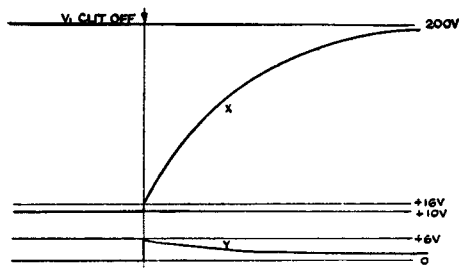


Fig. 105.—Waveforms of circuits of figs. 103 and 104

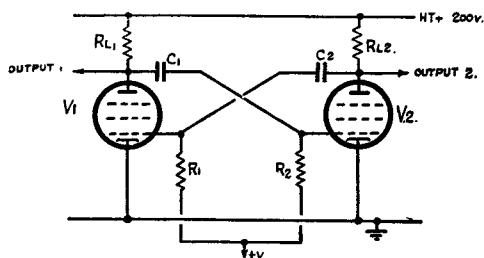


Fig. 106.—Skeleton circuit of pentode multivibrator

208. Fig. 106 shows the skeleton circuit diagram of a typical grid-coupled multivibrator. The grid resistors are returned to a positive potential which may be derived from a potentiometer chain across the HT supply.

Pentode valves have been chosen with the anode loads sufficiently large to produce bottoming at zero grid volts and under these conditions the waveforms of the circuit are as shown in fig. 107.

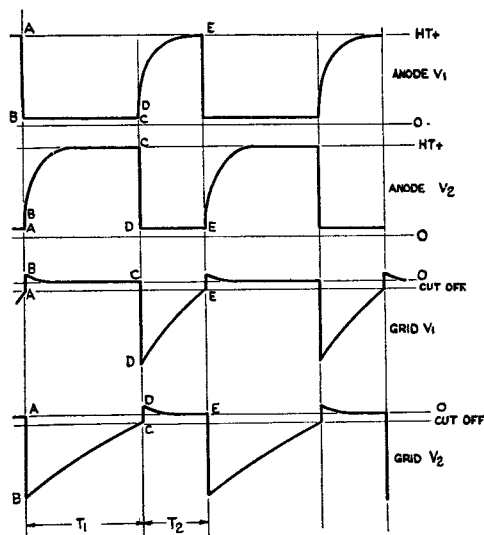


Fig. 107.—Waveforms of circuit of fig. 106

209. Due to the cross-coupling of anodes and grids the circuit is unstable when both valves are conducting. This can be seen as follows. Suppose the anode current of V_1 to fall slightly due to some cause such as random noise fluctuations. The anode voltage of V_1 will therefore rise and with it the grid voltage of V_2 . This change is amplified by V_2 and passed back as a fall to the grid of V_1 thus assisting the original anode current change. The valves are, in effect, acting as amplifiers, and due to the high gain the action is violently cumulative, the grid of V_1 being driven almost instantaneously beyond cut-off.

210. We are now in a position to explain the waveforms of fig. 107.

(i) *A to B.* The grid of V_2 is at zero and V_2 therefore bottomed while the grid of V_1 is just rising beyond cut-off. The p.d. across C_2 is therefore some 15 volts and remains at this value during the very rapid cumulative change. The anode of V_1 is initially at $HT+$ and the p.d. across C_1 is therefore 200 volts and this, too, remains constant.

(ii) *At B.* (a) The anode of V_1 is at the bottoming voltage of +10 volts, and since the p.d. across C_1 is 200 volts the grid potential of V_2 must be -190 volts.

(b) V2 has been cut off suddenly, and since the p.d. across C2 is 15 volts, the remaining 185 volts must be shared between R_{L2} and the grid-cathode path of V1. (Compare with figs. 103 and 104.) The grid voltage of V1 is therefore about +5 volts and the anode voltage of V2 about +20 volts, these being calculated for an anode load R_{L2} of 33K.

(iii) *B to C.* (a) The anode of V2 rises towards HT+ and the grid of V1 falls to zero as C2 charges through R_{L2} and the grid-cathode path of V1, the time constant being C2 R_{L2} approximately. (Compare with fig. 105.)

(b) Since the anode of V1 is bottomed, its potential remains steady at about +10 volts.

(c) The grid of V2 rises towards +V on a time-constant C1 R2.

(iv) *At D.* The grid of V1 is at zero with V1 bottomed, while the grid of V2 is just rising beyond cut-off. We have therefore reached the initial conditions of A but with the valves interchanged.

211. The remainder of the cycle may easily be followed through, the events of the whole cycle being most conveniently set out in tabular form as given below:—

	Anode V1	Anode V2	Grid V1	Grid V2
A	+200v	+10v	-5v	0
B	+10v	+20v	+5v	-190v
BC	Steady at +10v	Rises to +200v on time-constant C2 R_{L2} then remains steady	Falls to zero on C2 R_{L2}	Rises on C1 R2 towards +V
C	+10v	+200v	0	Reaches cut-off at -5v, initiating cumulative change
D	+20v	+10v	-190v	+5v
DE	Rises to +200v on C1 R_{L1} then remains steady	Steady at +10v	Rises on C2 R1 towards +V	Falls to zero on C1 R_{L1}

At E, V1 cuts on and the cycle repeats.

212. The time T1 (fig. 107) is that required for the grid of V2 to rise from -190v. to cut-off. As the grid is rising towards +V on the time constant C1 R2, a variation in T1 may be obtained either by varying R2 or by varying V. This is illustrated in fig. 108, where curve I is taken as the standard. Decreasing R gives curve II, while decreasing V gives curve III. A similar argument applies to T2.

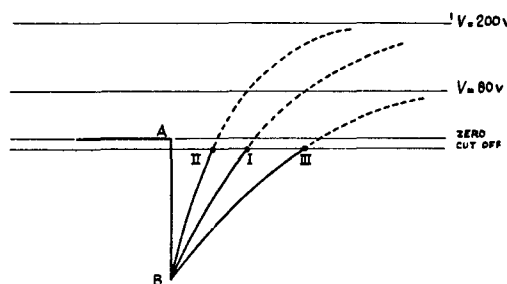


Fig. 108.—Variation of T1 of fig. 107 by two methods

213. The period of the multivibrator is $T1 + T2$, and if it is desired to alter this without changing the ratio $T1/T2$ the resistors R1 and R2 must either be made variable and ganged or else R1 and R2 must be kept fixed and V varied, the latter being preferable.

214. It will be noted that the waveforms at the anodes are not perfectly square. The negative-going edges, which are due to the cumulative changes, are extremely fast, but the positive-going edges are rounded due to the necessity of charging the coupling capacitors

through the anode loads. For good squareness these capacitors should therefore be kept small, but for a given V the period depends on the time constants C1 R2 and C2 R1, and this imposes a compromise design. If in an endeavour to keep C1 and C2 small, R1 and R2 are made too large, hum may be picked up on the grids, causing a jittery output waveform. A high value of V assists in the design and multivibrator circuits often have the grid resistors returned to a potential at or near the HT+ line.

215. If the anode loads are made too low for bottoming at zero grid volts or if triode valves are used the waveforms become more complicated. Fig. 109 shows how the waveforms are modified if R_{L1} is too low or if V1 is replaced by a triode. During the portion BC of the waveform there is a positive pip BB' on the grid waveform of V1, and since V1 is now no longer bottomed this appears as a negative pip on the anode waveform, the anode voltage being well above the bottoming value along B'C. The rise BB' is transferred via C1 to the grid of V2, hence the waveform at this grid is no longer a simple exponential along BC. The period T1 is therefore reduced and it is now impossible to calculate the period of the multivibrator from the circuit values.

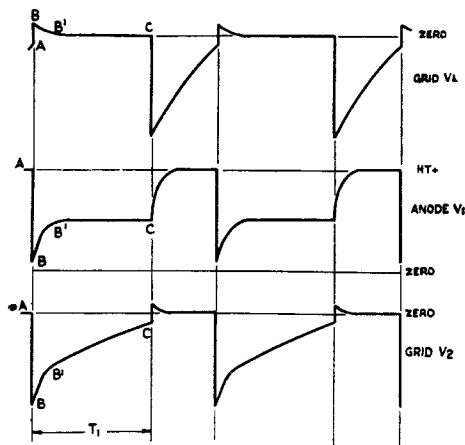


Fig. 109.—Waveforms of circuit of fig. 106 when V1 is not bottomed

The cathode coupled multivibrator

216. The cathode coupled multivibrator is used when a very asymmetrical waveform is desired, e.g. a pulse of 20 microsecs. duration occurring every 1000 micro-seconds. It has the following advantages:—

(i) Both edges of the pulse are produced by cumulative changes, thus very fast edges are possible.

(ii) Both positive and negative pulse outputs are available.

(iii) The pulse recurrence frequency may be varied without changing the pulse length.

(iv) The outputs from the circuit are derived from low impedance points. A typical circuit is shown in fig. 110 and the waveforms produced in fig. 111. The outputs are taken from the anode of V2 and the two cathodes, and since R3, R4 and R5 must be kept small

in order to give low output impedances, V2 must be capable of passing a heavy current during the pulse if a reasonable output amplitude is required. A valve of the VT60A type is therefore used for V2. V1 is a VR91.

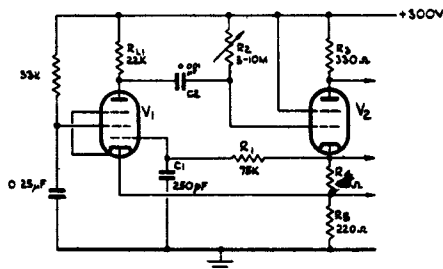


Fig. 110.—Circuit of cathode-coupled multivibrator

217. Starting at A, V1 is conducting normally (R_{L1} being chosen so that the valve is not bottomed), while V2 is cut off but just coming into conduction. We may then follow the cycle of operation:—

(i) *A to B.* V2 starts conducting and its cathode rises, causing the cathode of V1 to rise also. The grid potential of V1 is held at zero by the capacitor C1, the anode current of V1 therefore falls with a consequent rise in anode voltage which pulls up the grid of V2 with it. *Cumulative action.*

(ii) *At B.* V1 is cut off by the rise in cathode voltage and V2 is hard on. In the grid-coupled multivibrator the rise in anode voltage of the valve which is suddenly cut off is limited to about 15 volts, but in the circuit under consideration the rise is much greater.

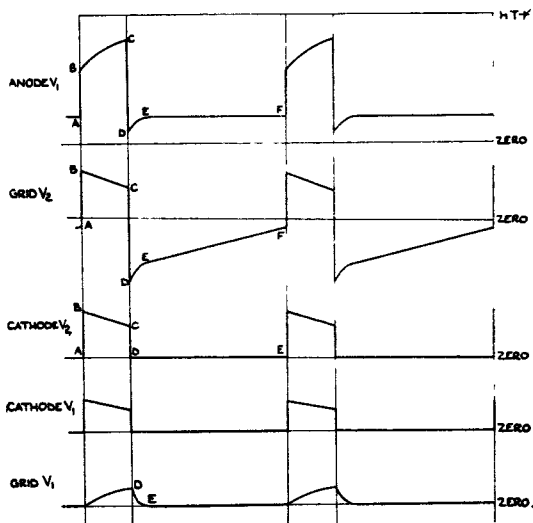


Fig. 111.—Waveforms of circuit of fig. 110

This is due to cathode follower action in V2 which allows the grid of V2 (and with it the cathode) to rise considerably before the grid current starts. At B the cathode of V2 is at about +60 volts with the grid slightly higher, the cathode of V1 is therefore at about +40 volts.

(iii) *B to C.* (a) C2 is charging up and the anode of V1 is rising towards HT+ while the grid of V2 is falling, carrying the cathode with it. (Cathode following.)

(b) C1 is charging up through R1 and the grid of V1 is therefore rising towards the cathode of V2 (about +60v) on a time constant C1 R1.

(iv) *C to D.* When the grid of V1 has risen sufficiently for the valve to start conducting the start of anode current initiates cumulative action, resulting in V2 being cut off.

(v) *D to E.* When V2 is cut off at D, the cathodes of V1 and V2 fall to zero (neglecting the small drop of voltage across R5 due to the current in V1). C1 had charged up to nearly +40 volts and now discharges rapidly through the grid-cathode path of V1. At D the grid of V1 was positive with respect to its cathode and the rapid fall of grid voltage as C1 discharges explains the pips on the waveforms of the anode of V1 and the grid of V2.

(vi) *E to F.* The grid of V2 rises towards HT+ on a time constant C2 R2 until V2 cuts on and the cycle repeats.

218. *Note.*—(i) The pulse length is determined by C1 and R1, while the recurrence period is determined by C2 and R2.

(ii) Positive pulse outputs are obtained from the cathodes; the load in the anode of V2 enables a negative pulse to be taken out at this point.

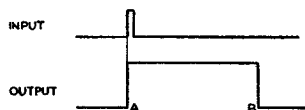


Fig. 112.—Illustrations of definition of "delay circuits"

(iii) Since bottoming is not employed, variations in HT supply voltage and changing of valves affect the output waveforms and the recurrence frequency more than in the grid-coupled circuit.

(iv) While the edges of the pulse are very fast, the top is not flat.

Delay circuits

219. Consider fig. 112. It is required to produce a square waveform AB from a short input pulse, the front edge A of the output coinciding with the triggering pulse and the duration AB being variable. Two circuits to accomplish this will be considered:—

- (i) The flip-flop (using two valves).
- (ii) The single-valve delay circuit.

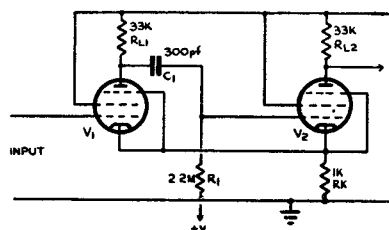


Fig. 113.—Circuit of flip-flop

The flip-flop

220. The circuit of a simple type of flip-flop is given in fig. 113, and the waveforms in fig. 114. The width of the triggering pulse is deliberately exaggerated.

221. Until the triggering pulse arrives, the grid of V1 is at zero and the circuit in the stable condition. V2 is conducting with its grid held at cathode potential by grid current, and both grid and cathode are therefore at some 10 to 15 volts above zero. The cathode of V1 is at the same level and V1 is thus cut off. The anode of V1 is at HT+ and V2 is bottomed.

(i) *A to B.* The leading edge of the triggering pulse drives V1 into conduction and the valve bottoms. The grid of V2 falls by an equal amount, cutting V2 off and anode of V2 rises to HT+. The cathode potential changes slightly as the current changes over from V2 to V1, the amount of the change depending on the amplitude of the triggering pulse.

(ii) *At C.* Since V2 is cut off we may neglect it temporarily and consider V1 only. When the grid of V1 returns to zero at the end of the triggering pulse, the anode current falls to a value determined by the valve characteristics, supply voltages and the cathode bias resistor RK. With the values shown in fig. 18 the valve will be biased well back with the cathode at about +3 volts. With the given anode load RL1 the anode will rise to about $\frac{1}{2}$ HT and the grid of V2 will rise by an equal amount, but V2 will still be cut off.

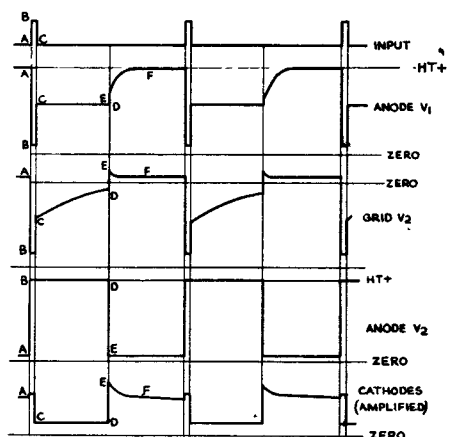


Fig. 114.—Waveforms of circuit of fig. 113

(iii) *C to D.* The anode of V1 remains at $\frac{1}{2}$ HT and C1 discharges, the grid of V2 rising towards $+V$ on a time constant $C1 R1$.

(iv) *D to E.* At D the grid of V2 has risen far enough for the valve to come into conduction. The flow of current in V2 raises the cathode potential and decreases the current in V1 with a consequent rise in anode potential. This rise is passed on to the grid of V2, increasing the current in that valve still further. Cumulative action thus takes place and V1 is cut off by the sharp rise in cathode potential. At E, V2 is bottomed with its grid slightly positive w.r.t. the cathode.

(v) *E to F.* C1 now charges through R_{L1} and the grid-cathode path of V2 on a time constant approximately equal to $C1 R_{L1}$. The anode of V1 moves towards HT+ and the grid of V2 falls, the cathode falling with it by cathode follower action. At F the original stable state has been restored.

222. *Notes.*—(i) The “hold over” or delay time (period when V2 is not conducting) may be controlled either by variation of $R1$ or V , the potential to which $R1$ is returned.

(ii) Some care must be exercised in the choice of R_K . If R_K is too low, V1 cannot be cut off by the flow of current in V2. No stable state is then possible as the circuit multivibrates. If R_K is too large, the anode current of V1 will be too small during the hold-over time; the anode of V1 will thus rise too far at the point C, and in the extreme case the grid of V2 may be brought above cut-off at this point, the circuit then refusing to hold over at all.

(iii) If R_K is correctly chosen but R_{L1} is too small, similar effects are produced. The voltage drop across R_{L1} is insufficient so that

the anode of V1 will rise to a potential well above $\frac{1}{2}$ HT at the point C, bringing the grid of V2 up to a point near to or above cut-off.

(iv) If R_{L2} is too small, V2 will not be bottomed at the point E and a negative pip will appear on the anode waveform (see grid-coupled multivibrator). This will not affect the operation of the circuit, which is independent of R_{L2} .

(v) It has been assumed that a very narrow positive pulse is used for triggering. If it is necessary to use a wider pulse it is preferable to differentiate this before applying it to the grid of V1.

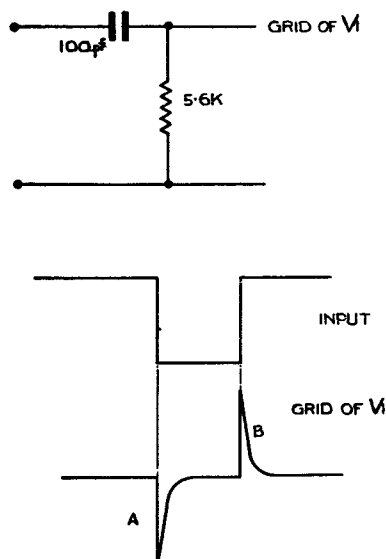


Fig. 115.—Triggering of flip-flop from wide negative-going pulse

(vi) *Negative pulse* (fig. 115). The negative pip A will not affect the circuit as it occurs during the stable state when V1 is already cut off. The positive pip B then triggers the circuit at the back edge of the input pulse.

(vii) *Positive pulse.* The pips A and B are now of reversed polarity and the circuit is triggered by A at the front edge of the input pulse. The negative pip at the back edge must be removed since a negative pulse applied to the grid of V1 during the hold-over period would cut that valve off and initiate prematurely the cumulative change at D. A diode must therefore be used as in fig. 116 to remove the negative pip.

The single-valve delay circuit

223. The circuit is shown in fig. 117 and the waveforms in fig. 118. The input pulse

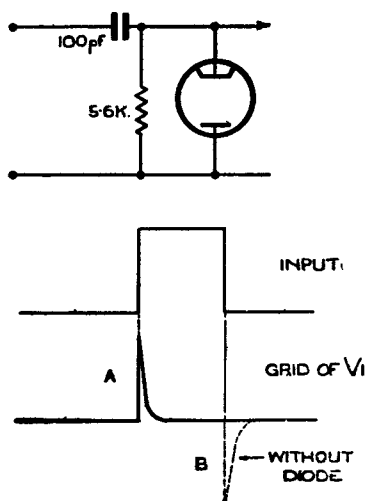


Fig. 116.—Triggering of flip-flop from wide positive-going pulse

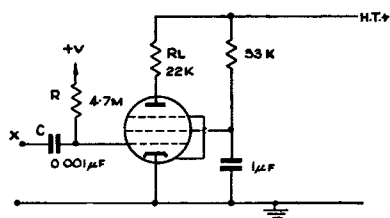


Fig. 117.—Single-valve delay circuit

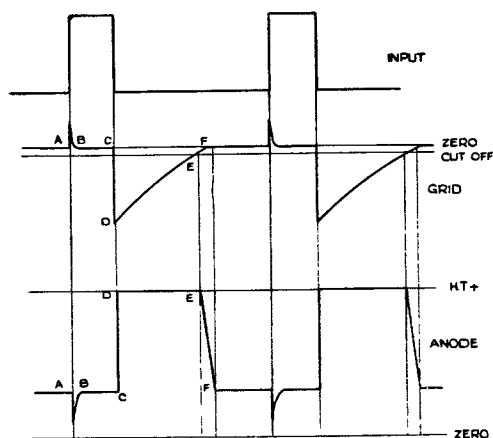


Fig. 118.—Waveforms of circuit of fig. 117

should be of positive polarity, square shape, of at least 20 microseconds duration and of about 100 volts amplitude. The width has been deliberately exaggerated in the figure. The initial stable state of the circuit is with the grid held at zero by grid current and the anode at a fairly low potential but not bottomed.

(i) *A to B.* The input pulse raises the potential of the point X to +100 volts and C charges up very rapidly through the grid-cathode path of the valve. A positive pip therefore occurs on the grid waveform and a corresponding negative pip on the anode waveform. The exact size of the positive pip at the grid cannot be given as it depends on the output impedance of the generator producing the input pulse. At B the grid has returned to zero and C is therefore charged to +100 volts.

(ii) *C to D.* At the back edge of the input pulse the potential of X falls to zero and the grid drops by an equal amount, i.e. to -100 volts. The valve is cut off and the anode rises to HT+.

(iii) *D to E.* C now discharges, the grid moving towards +V on a time constant CR.

(iv) *E to F.* At E the grid reaches cut-off and the flow of anode current causes a progressive fall in anode potential as the grid rises towards zero along EF. At F the grid is held by grid current and the initial conditions are restored.

224. *Notes*—(i) The delay time may be controlled by varying either R or V, and depends also on the amplitude of the input pulse.

(ii) The back edge EF of the output (anode) waveform is much less steep than in the flip-flop as it is produced not by a cumulative change bringing the valve into conduction very quickly, but by the relatively slow rise of grid voltage along EF. The steepness is reduced as the delay is increased.

The blocking oscillator

225. The blocking oscillator is used for the direct generation of fairly short pulses, regenerative coupling between different valve electrodes being used to give a rapid switching action which produces reasonably steep edges in the pulse. In the circuit of fig. 119 the coupling is between grid and screen, but coupling between grid and cathode circuits may also be used. The iron-cored transformer has a turns ratio n of about 3, the grid winding

being tuned by the capacitance C_1 which is usually just the self-capacitance of the winding. The waveforms are given in fig. 120.

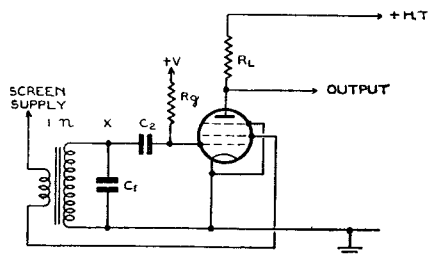


Fig. 119.—Circuit of typical blocking oscillator

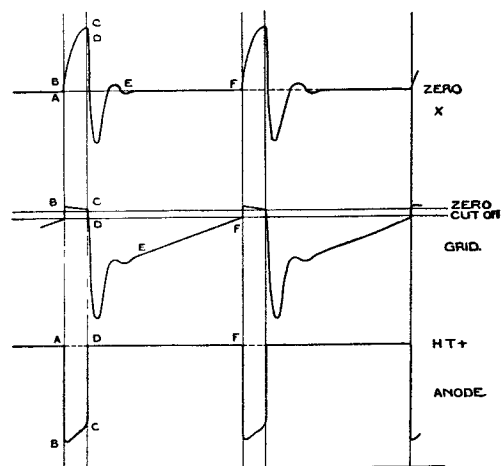


Fig. 120.—Waveforms of circuit of fig. 119

(i) *A to B.* The valve is just coming into conduction and the anode and screen currents are rising. The change in screen current induces an e.m.f. in the LC_1 circuit, causing a current to flow which charges C_1 , raising the grid potential. The rise in screen current is thus accelerated, the close coupling between screen and grid coils producing a very strong cumulative effect which moves the grid potential across the grid base at a very fast rate. The anode current rises also and the anode potential falls to a value at or near bottoming, depending on the load resistance R_L .

(ii) *B to C.* When the grid potential reaches zero, grid currents start and the current flowing in L must now charge C_2 as well as C_1 . The grid potential is now not the same as that of the point X , but is determined solely by the voltage drop across the grid-cathode path of the valve produced by the charging current of C_2 . It is this charging current which causes the grid to be slightly positive at B . During BC the grid potential is falling, but is so nearly constant

that the resultant variations in screen current produce negligible induced voltages in the transformer windings. (A more detailed mathematical treatment shows that, provided the grid base of the valve is small, if the screen current at B is I_s , the current flowing in L is $\frac{I_s}{n}$. This current decreases as C_1 and C_2 charge until at C , when the potential of X has reached its maximum, it has fallen to zero. The portion BC of the waveform at X is one-quarter of a cycle of oscillation of the circuit made up of L shunted by C_1 and C_2 , this oscillation being "shocked" into existence by the sudden switching on of current in the screen circuit during AB .)

(iii) *C to D.* As the potential of X passes through its maximum the current in L must reverse, this current being produced by the discharge of C_1 . C_2 cannot discharge in this way since the grid-cathode path of the valve acts as a diode preventing current reversal. As the potential of X falls, that of the grid therefore falls with it, the cumulative effect now taking place in the reverse direction until the valve cuts off at D .

(iv) *D to E.* The sudden switching-off of the screen current produces a "shocked" oscillation in the LC_1 circuit, which is rapidly damped out. At the same time C_2 begins to discharge through R_g .

(v) *E to F.* C_2 continues to discharge through R_g , the grid potential rising towards $+V$ on a time constant $C_2 R_g$ until it reaches cut-off at F and the cycle repeats.

226. *Notes.*—(i) The pulse length is determined by the period of oscillation in the L, C_1, C_2 circuit, i.e. by the product LC_2 , since C_2 is usually much greater than C_1 .

(ii) The maximum positive excursion of X (point C) depends on C_2 , the screen current at B and the transformer ratio n . If the valve is bottomed at B , more current flows to the screen and X is driven further positive.

(iii) The damping of the LC_1 oscillation during DE must be heavy. With a low degree of damping there is a danger that, after one cycle of this oscillation, the grid potential may again be carried above cut-off, thus producing a double pulse or even continuous oscillation. The transformer is often designed so that resistance losses in L and iron losses in the core are sufficient to provide the requisite damping. Alternatively, additional damping may be provided by connecting a resistance in shunt with the grid coil. The oscillation during DE is not essential to the operation of

the circuit, which will work equally well if the damping is greater than the critical value.

(iv) The recurrence period depends on (a) the time constant $C2 R_g$, (b) the potential $+V$ to which R_g is returned, (c) the extent to which $C2$ is charged during the pulse, i.e. the maximum positive excursion of X . It may most conveniently be varied by returning R_g to a potentiometer across the HT supply.

Production of pulses from square wave

227. Given a square waveform it is required to produce from it a series of pulses of a recurrence period equal to the period of the square wave. Three methods are commonly used:—

- (i) Short CR circuit
- (ii) Ringing circuit
- (iii) Delay line

and each of these is usually combined with a pip-eliminating and pulse-shaping valve or valves.

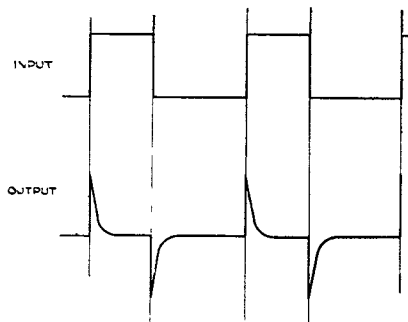


Fig. 121.—Action of short CR circuit

The short CR circuit

228. The input square wave and the associated output pips from the short CR or differentiating circuit are shown in fig. 121. The objections to using this output directly are:—

- (i) Both positive and negative pips are present.
- (ii) The pips are of poor shape, having a steep front edge but an exponential back edge, and thus being very different from the ideal square pulse. The pip eliminator valve may be set to remove, say, the negative pips and at the same time it will much improve the shape of the positives.

229. The basic circuit is shown in fig. 122 and the close resemblance to the squarer circuit may be seen by comparing this with fig. 102.

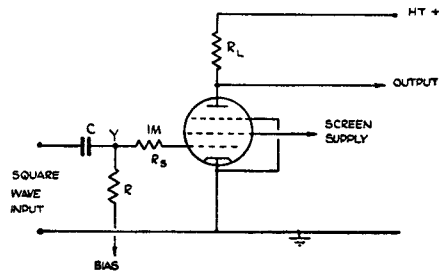


Fig. 122.—Basic circuit of "pip" eliminator stages

This resemblance is further emphasised by lettering fig. 121 and 122 to correspond with figs. 100 and 101. The case of negative bias is shown in fig. 123. The valve is cut off except during a short portion of the input positive pip, and thus only this pip has been shown in the figure.

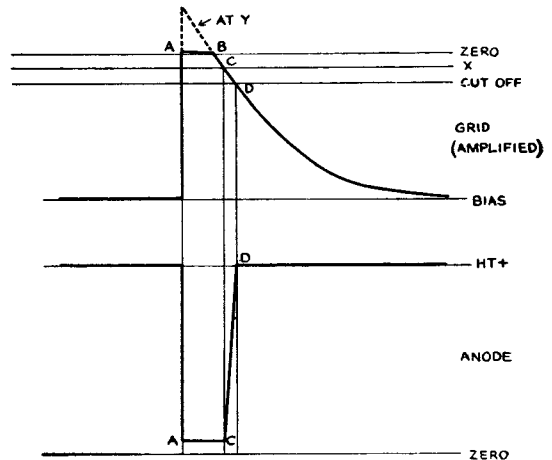


Fig. 123.—Waveforms of circuit of fig. 122 with negative bias

(i) *At A.* The point Y has been carried positive by the positive-going edge of the square wave but the grid is held just above zero by grid current in conjunction with the stopper R_g . The anode falls very quickly from HT+ to the bottoming voltage.

(ii) *A to B.* The grid falls very slowly during the exponential back edge of the pip and the anode remains bottomed.

(iii) *B to C.* At B the grid ceases to be held by grid current and therefore falls with Y, but the anode is still bottomed until the point C is reached.

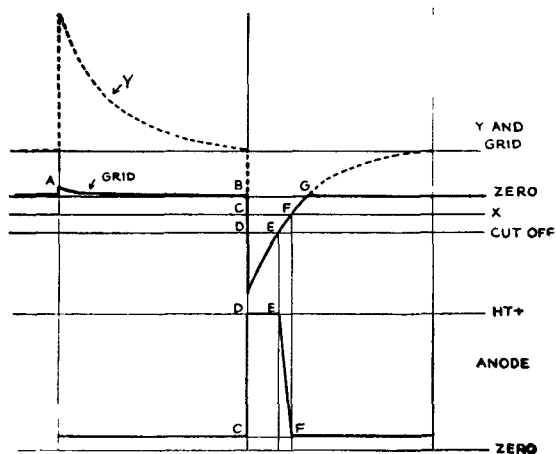


Fig. 124.—Waveforms of circuit of fig. 122 with positive bias

(iv) *C to D.* The grid falls rapidly through the grid base until the valve cuts off at D, the anode meanwhile rising from the bottoming voltage to HT+.

The corresponding case of positive bias is shown in fig. 124. With no input the grid is held at about zero by grid current.

(v) *A to B.* The positive pip at Y causes a very small positive movement of the grid, which does not affect the anode since the latter is bottomed.

(vi) *B to D.* The front edge of the negative pip drives the grid beyond cut-off and the anode rises to HT+.

(vii) *D to E.* The valve remains cut off.

(viii) *E to F.* The grid potential is rising and the anode falls from HT+ at E until bottoming occurs at F.

(ix) *F to G.* The further rise in the grid potential does not affect the anode.

230. *Notes.*—(i) The front edge of the output pulse is very steep, the back edge less so.

(ii) Provided the valve is well bottomed at zero grid volts, the output at the anode is an amplified and inverted copy of the input waveform included between the “cut-off” and “X” levels.

(iii) With *negative* bias the *positive* input pip is selected and gives a *negative* output pulse. For *positive* bias the output pulse is *positive*.

(iv) The width of the output pulse may be controlled either by variation of bias or by variation of R in the differentiating circuit.

(v) The practical circuit diagram giving a choice of positive or negative bias is the same as fig. 102, the values of C and R being chosen to give the required differentiation.

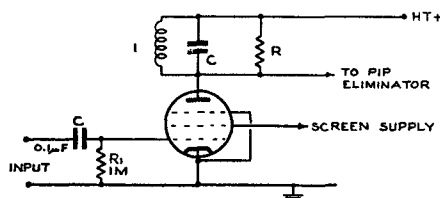


Fig. 125.—Circuit for production of pulses by ringing circuit

The ringing circuit

231. The circuit is shown in fig. 125, and the waveforms in fig. 126. The square wave at the grid has its top clamped to the zero level by D.C. restoration, the grid-cathode path of the valve acting as a diode. A damped oscillation will be produced at the anode every time the valve is turned on or off. If the damping is correctly chosen the initial half-cycle will be of much larger amplitude than the succeeding half-cycles and the output at the anode will approximate to a series of pips, alternately positive and negative. These pips may be dealt with in the usual way by a pip eliminator valve.

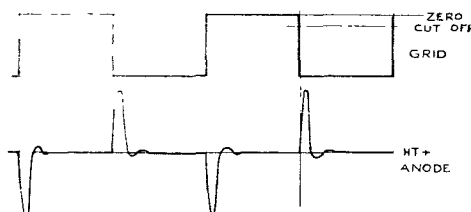


Fig. 126.—Waveforms of circuit of fig. 125

Notes.—(i) The shape of the damped half sine wave pips is better than the differentiated pips produced by a short CR circuit, but an extra valve is needed.

(ii) The damping needs fairly careful adjustment; if too heavy, the pips will be of small amplitude, if too light more than one half cycle may pass through the pip eliminator valve.

(iii) The time constant of the coupling to the pip eliminator valve should be long enough to avoid differentiation of the output of the ringing circuit.

The delay line

232. The usual type of transmission line (e.g. coaxial) has distributed inductance and capacitance. If L = inductance per cm

length, C = capacitance per cm length, then an impulse will be propagated along the line

with a velocity of $\frac{1}{\sqrt{LC}}$ cms per second.

With a coaxial line with polythene dielectric this velocity is about $\frac{2}{3}$ of that of electromagnetic waves in free space, i.e. about 2×10^{10} cms per second. It is impracticable to use such a line for any but the very shortest delays, a simple calculation showing that to produce a delay of 10 microseconds between sending and receiving ends, a line $1\frac{1}{4}$ miles long would be needed. In order to give delays of this order

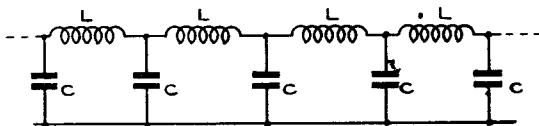


Fig. 127.—Delay line showing four sections

an “artificial transmission line” or “delay line” is used. In this the inductance and capacitance are “lumped” instead of being distributed (fig. 127 and 128). With a delay line

if L = inductance per section in henries and C = capacitance per section in farads

the velocity of propagation is $\frac{1}{\sqrt{LC}}$ sections per second and the characteristic impedance

$$Z_0 = \sqrt{\frac{L}{C}}$$

Thus if T_0 = required time delay in seconds

and n = number of sections

$$\text{then } T_0 = n \sqrt{LC} = \sqrt{nL \cdot nC}$$

$$= \sqrt{\text{Total } L \times \text{Total } C}$$

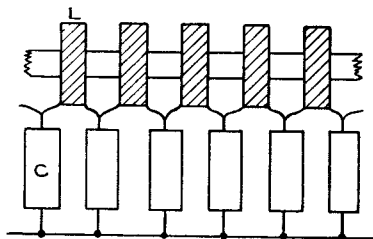


Fig. 128.—Five sections of delay line showing construction

are equivalent to distributed inductance and capacitance to a first approximation only. Such distortion is neglected in what follows and it is assumed that the delay line may be treated as a transmission line. Consider the circuit of fig. 129. D.C. restoration takes place at the grid (compare fig. 125) and thus the valve is turned on suddenly at the positive-going edge of the square wave and turned off suddenly at the negative-going edge. There are three possible cases; R equal to, less than and greater than Z_0 . The line is assumed to be short-circuited at the end remote from the anode.

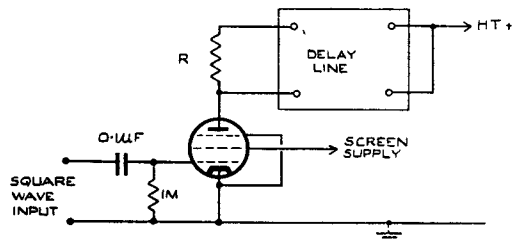


Fig. 129.—Circuit for production of pulses by delay line

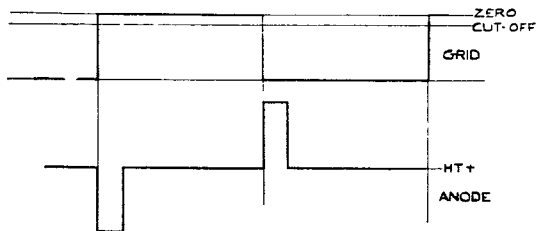


Fig. 130.—Waveforms of circuit of fig. 129, when $R = Z_0$

(i) $R = Z_0$.—When the valve is turned on, the load in the anode circuit is R in parallel with the Z_0 of the line, i.e. $\frac{1}{2} Z_0$. A voltage drop of $\frac{1}{2} Z_0 I$ thus takes place at the anode and this voltage step is propagated down the line until it reaches the far end. Since this end is short-circuited the negative voltage step is reflected as an equal positive step which travels back along the line to the sending end. On arrival it cancels the voltage drop already present and since the line is terminated with Z_0 at the valve end no reflection of the positive step takes place. Hence a single negative-going pulse is obtained at the anode, the width of this pulse being $2T_0$, the time for the voltage step to make the double journey along the line. Similarly, when the valve is turned

off a single positive-going pulse is obtained at the anode (fig. 130).

(ii) $R < Z_0$.—When the valve is turned on the drop in voltage at the anode is $\frac{RZ_0}{R + Z_0} I$ and this step is propagated down the line and reflected with a reversal in sign as in the previous case. When the reflected (positive) voltage step arrives at the anode the line is not now terminated with Z_0 but with a resistance lower than this value and so the reflected positive step not only cancels the original voltage drop but is itself partially reflected with a change of sign. Thus at the end of the time $2T_0$ the voltage at the anode does not return to its original value (HT+) but to a lower value. The process continues with successive reflections of decreasing amplitude until the voltage changes are no longer detectable. Similar effects occur when the valve is turned off (fig. 131).

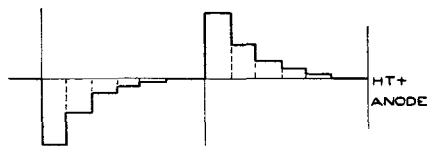


Fig. 131.—Waveforms of circuit of fig. 129, when $R < Z_0$

(iii) $R > Z_0$.—In this case the positive voltage step reflected from the far end of the line is reflected at the valve end without a change of sign and the resultant waveform is shown in fig. 132.

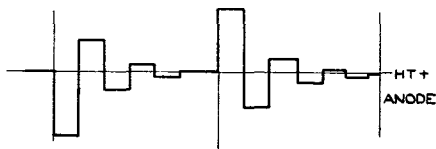


Fig. 132.—Waveforms of circuit of fig. 129, when $R > Z_0$

234. Notes.—(i) The waveforms of figs. 130 to 132 have been drawn neglecting the distortion which inevitably occurs in a delay line and which is most marked with a number of successive reflections.

(ii) As single pulses are needed in practice, the line is operated in the matched condition ($R = Z_0$).

(iii) The positive or negative pulses in the output waveform may be removed by a pip eliminator stage.

(iv) Owing to the square shape of the pulse, variation in bias of the pip eliminator valve cannot be used as a pulse width control. The width can only be varied by altering the number of sections and this is usually carried out by a switch which moves the short circuit termination along the line.

Production of a delayed pulse by means of a delay line

235. The circuit of fig. 129 may, with slight modifications be used to produce a delayed pulse, the modified circuit being shown in fig. 133 and the normal waveforms in fig. 134. Comparison of figs. 129 and 133 shows that the resistor R has been moved from the valve end of the delay line to the far end where it replaces the short-circuit termination. The output is also taken from the far end instead of from the anode of the valve.

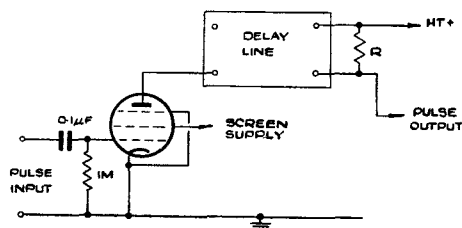


Fig. 133.—Circuit for production of delayed pulse

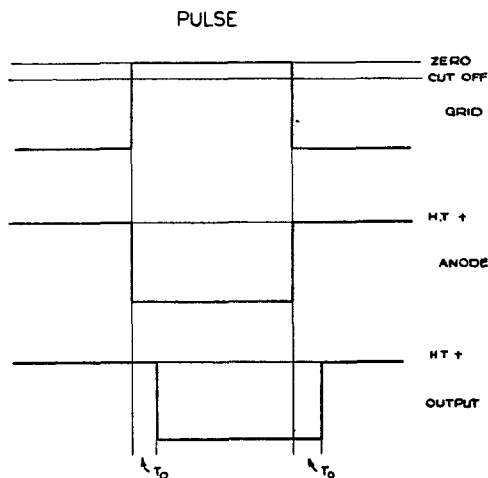


Fig. 134.—Waveforms of circuit of fig. 133 with $R = Z_0$

236. With $R = Z_0$, which is the normal working condition, it may easily be seen that the anode load of the valve is a length of transmission line terminated in its characteristic impedance, i.e. the valve is working into

a pure resistance of value Z_0 . When the valve is turned on at the front edge of the input pulse the anode current rises suddenly from zero to I , the anode voltage therefore falling by an amount $Z_0 I$, and this negative step is propagated down the line, reaching the far end after a time T_0 . Since the line is correctly terminated no reflection takes place so that the potential drop is maintained until the end of the input pulse. When the valve is cut off the positive step at the anode is again propagated along the line, cancelling the initial negative voltage as it goes. The waveform at the far end of the line is thus the same as that at the anode but is delayed by T_0 (fig. 134).

The terminating resistor R must be kept in position at the far end of the line and not moved with the output tapping.

CATHODE RAY TUBES

Principles and functions

239. A cathode-ray tube, or CRT for short, is a vacuum valve by means of which a spot of light can be produced on a screen at one end of the valve. The position and intensity of the spot can be altered almost instantaneously at will to give an indication of the position; i.e. range, azimuth and elevation, of any target as determined by a radar system. This property makes the CRT the most usual



Fig. 135.—Output waveform of circuit of fig. 131 when $R \neq Z_0$

237. When the terminating resistor R is not equal to the characteristic impedance Z_0 of the delay line, partial reflection of a voltage step occurs, the reflected step travelling back along the line and being totally reflected at the valve end which, for a pentode valve, may be considered to be open-circuited. Successive reflections of decreasing amplitude therefore occur in much the same way as in the circuit of fig. 129, typical waveforms being shown in fig. 135.

238. *Notes.*—(i) The delay produced by the circuit of fig. 133 is T_0 , i.e. half the length of the pulse produced by the circuit of fig. 129.

(ii) In order to vary the delay time, the output is taken from a tapping on the line, the position of which may be varied by a switch.

indicating device in such a system, since very small intervals of time have to be measured.

240. The basic construction of a CRT is shown in fig. 136. The spot of light is produced by means of a beam of electrons from an "electron gun" G , this beam being focussed on a glass screen as shown. A device for deflecting the beam to any spot on the screen is inserted immediately after the gun. On the inner surface of the screen is a layer of fluorescent powder which glows when struck by the beam, thus producing the spot of light. Many types of screen are used, depending on the type of display required. The two main types of screen are (i) persistent, in which the screen continues to glow after the beam has been removed, and (ii) non-

persistent, in which the screen ceases to glow immediately the beam is removed. Persistent screens are usually used in PPI displays and non-persistent in range-amplitude displays.

241. The focussing and deflection of the electron-beam may each be done either electrostatically or electromagnetically; the usual combinations of these systems found in practice are:—

- (i) Electrostatic focussing and deflection.
- (ii) Electromagnetic focussing and deflection.
- (iii) Electrostatic focussing and electromagnetic deflection.

These three methods will be discussed separately.

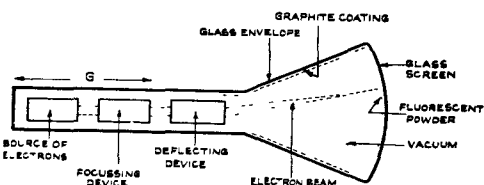


Fig. 136.—Diagram of the elements of a CRT

Electrostatic focussing and deflection

Description

242. The “gun” and deflector system of a CRT of this type is shown in fig. 137. A cathode coated with rare-earth oxides, is heated as in a normal valve and emits electrons. The flow of electrons and hence the intensity of the beam and resulting light-spot is controlled by the potential, relative to cathode, of a “grid”. This takes the form of a cylinder—sometimes known as the Wehnelt cylinder—with an aperture through which the electrons pass. The electrons are then accelerated by

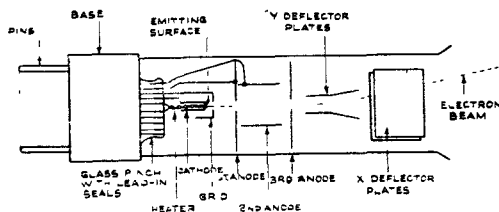


Fig. 137.—Gun and deflector plates of an electrostatically-focussed and deflected CRT

positive potentials applied to three anodes (numbered consecutively away from the cathode). The first and third anodes take the form of plates, with planes perpendicular to the axis of the electron beam, and apertures in the centres through which the beam passes.

The second anode is a cylinder coaxial with the beam. Its potential can be varied at will and gives a focus control, as explained later. Beyond the third anode is placed the deflector system, which consists of two pairs of plates at right-angles, to which the signals to be examined are applied. The resulting electric fields between the plates deflect the beam in two directions at right-angles, known as the “X” and “Y” directions by analogy with the horizontal (X) and vertical (Y) axes used in drawing graphs.

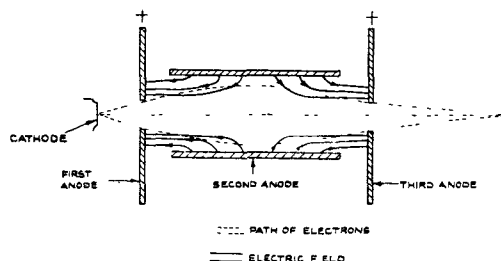


Fig. 138.—Electrostatic focussing

Functions of the electrodes

243. (i) *Cathode and grid.*—These perform functions similar to the counterparts in a normal valve and need not be discussed further.

(ii) *Anode system.*—The anodes perform two functions; first they serve to accelerate the electrons in the beam towards the screen, and second they focus the beam on the screen. The second anode is held negative with respect to the other two, and the resulting field pattern is shown in fig. 138. The arrows indicate the direction in which the electrons are urged, and not the conventional direction of the field. It can be seen that electrons which lie on the axis of the system do not undergo any deflection, while those that diverge from the axis are directed back towards it, giving a focussing action. The potentials of the first and third anodes are fixed, and that of the second anode can be varied at will, altering the field pattern in the system and hence providing a means of controlling the focus.

244. The fineness of focus is limited by the mutual repulsion of the electrons in the beam; the smallest spot obtainable is usually about 0.5 mm. in diameter.

Deflector plates

245. These are shown in perspective in fig. 139(a), and one pair of plates in section in fig. 139(b), (c). Suppose that a signal voltage is applied to the plates so that at a

given instant the top plate is at a positive potential with respect to the bottom one. Then the electric field between the plates will accelerate the electrons by an amount proportional to the field in the direction shown and the beam will be deflected upwards. If the sign of the applied voltage is changed, the beam will be deflected downwards. The deflection of an electron is proportional also to the time it is between the deflector plates, i.e. is inversely proportional to the velocity of the beam.

246. In fig. 139 (b) the maximum angle through which the beam can be deflected is θ ; this can be made larger by splaying out the ends of the plates nearer the screen, allowing the greater angle of deflection θ' . It will be seen that the beam is deflected in a plane perpendicular to the plane of the plates; consequently, the X plates, which deflect the beam horizontally, are vertical and similarly the Y plates are horizontal. The pair of plates nearer the anode system has a greater sensitivity than the other pair, and consequently these are usually made the Y plates; any time-base waveform which is applied to the X plates can easily be made of sufficient amplitude to compensate for the lower sensitivity of these plates.

The graphite coating

247. A coating of colloidal graphite is usually placed on the inner wall of the glass bulb, and is internally connected to the third anode (see fig. 136). This serves two purposes. The screen, due to the impact of the electron beam, not only glows but also emits "secondary" electrons of slow velocity. If these were not dispersed a heavy negative charge would build up on the screen which eventually would be sufficiently strong to repel the electron beam and prevent it reaching the screen, thus rendering the tube inoperative. The graphite coating, since it is connected to the third anode, attracts the secondary electrons and removes them. Secondly, the coating acts as an electrostatic screen and prevents stray external electrostatic fields from producing unwanted deflection of the electron beam.

248. Stray magnetic fields will also deflect the beam (see "electromagnetic deflection") and the effect of these is minimised by placing a screen of high permeability material ("mumetal") round the tube.

Power supplies for electrostatic tubes

249. The voltages required for a typical 4 kV tube (VCR.517) are:—

Cathode,—3.95 kV
Grid, about—4.00 kV } V_{gk} —50v, variable
for brilliance
control

Anode 1,—2.0 kV

Anode 2, about—3 kV, variable for focus control

Anode 3, Earth

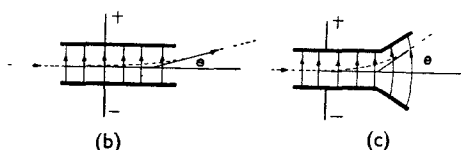
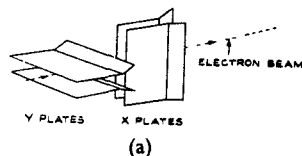


Fig. 139.—Deflector plates

250. The mean potential of the deflector plates must be that of the final anode so that distortion does not occur. Signals to the deflector plates are therefore applied through condensers and high resistance leaks are connected between the plates and the final anode. So that the condensers need not be of high working voltage it is usual to connect the final anode to earth and take the cathode to a high negative potential rather than earth the cathode and take the third anode to a positive potential. This also prevents spurious deflection of the spot by fields between the final anode and objects at earth potential near the screen.

251. A typical power supply circuit is shown in fig. 140. The high voltage from the secondary of transformer is rectified by V1 with its reservoir condenser C1, and smoothed by the filter R1 C2. Resistance smoothing is permissible, as the total current drain is only 1–2 mA, so that the voltage drop across R1 is small. The resulting H.T. is applied across the potentiometer chain P4, R6, P3, etc., which provides tapping points at suitable potentials to which the electrodes of the CRT V2 may be connected. The cathode current of V2 is of the order of 30–150 μ A so that a drain of 1–2 mA down the potentiometer chain is quite sufficient.

252. P4 varies V_{gk} and hence is the brilliance control. Blackout and brilliance signals are applied to the cathode and grid through the

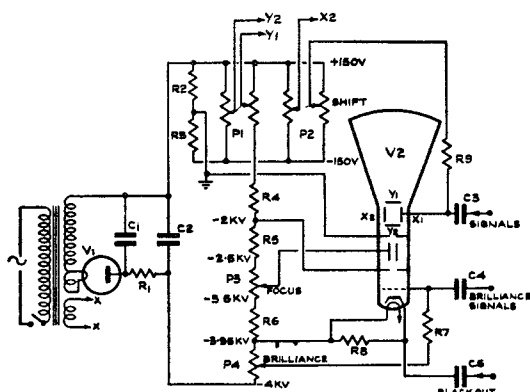


Fig. 140.—Power supplies for an electrostatic CRT

R-C couplings C5, R8 and C4, R7. It is important to note that C4 and C5 must withstand a high working voltage of the order of 4 kV. P3 varies with V_{a2} and hence is the focus control. The third anode is returned to earth and the deflector plates (of which only one is shown connected in the diagram) are fed through R-C couplings similar to C3 R9. C3 does not have to withstand more than 300–400 volts, and so may be of a normal type.

253. To provide a means of shifting the spot DC voltages are applied to the plates by returning the leaks R9 not to earth but to the sliders of potentiometers P1, P2 across which a potential of about 300 volts is applied. The mean of this potential is adjusted to be that of earth by resistances R2 and R3. The potentiometers P2 are ganged in opposition so that antiphase shift voltages are applied to the pair of deflector plates with which P2 is associated. This is necessary to avoid distortion.

254. The heaters of V2 are supplied from a separate heater winding XX on the transformer; both this and the rectifier heater winding must be well insulated to withstand the high working voltages. The heaters of V2 must be connected to a point near cathode potential, usually the “dead” side of the cathode leak R8.

Distortion in electrostatic CRTs

255. Distortion may arise from mechanical or electrical causes; the former gives rise to “astigmatism” and the latter to faults known as “deflection defocus” and “trapezium distortion”.

256. *Astigmatism.*—This is caused by misalignment of the anode system giving rise to faults in focus similar to astigmatism in

optical systems (fig. 141). The beam, instead of being focussed to a spot, has the form shown in the two sections. The picture on the screen, if this were placed at (a), is a vertical line of light, and at (b) a horizontal line. At some position (c) between (a) and (b) the beam produces an ill-defined circular patch of light, known as the “circle of least confusion”. Similarly, if the screen is considered fixed and the focus control of the tube is altered, the picture on the screen will pass from (a) through (c) to (b). It is usual to operate a tube suffering from astigmatism with the beam focussed on the circle of least confusion, though vertical lines in any diagram which may be depicted on the screen are more sharply defined (which is desirable in some cases) if the beam is focussed as in (a).

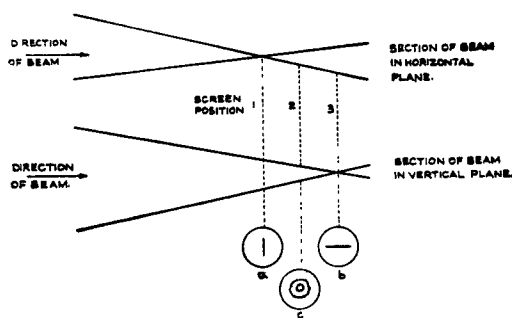


Fig. 141.—Astigmatism

257. The fault can sometimes be minimised by arranging that the mean potential of the deflector plates is not that of the third anode, i.e. $R2 \neq R3$ in fig. 140. This introduces unbalanced electric fields between plates and anode which have a focussing action on the beam and thus may correct for astigmatism.

Deflection defocus

258. This arises when unbalanced deflector voltages are used, i.e. one of a pair of deflector plates is connected to the third anode and signals are applied to the other (fig. 142). Consider the case when the top plate is returned to a positive potential +2V. Then, in addition to the deflecting field between the plates, there will be an asymmetrical field between the positive plate and third anode. This will distort and defocus the beam. Such a distorting field will be produced when the potential of the top plate is removed from that of the third anode, either positively or negatively. If, therefore, any waveform is applied to this deflector plate, and the beam is focussed at the centre of the tube, i.e. when there is no deflecting voltage, the spot will be enlarged

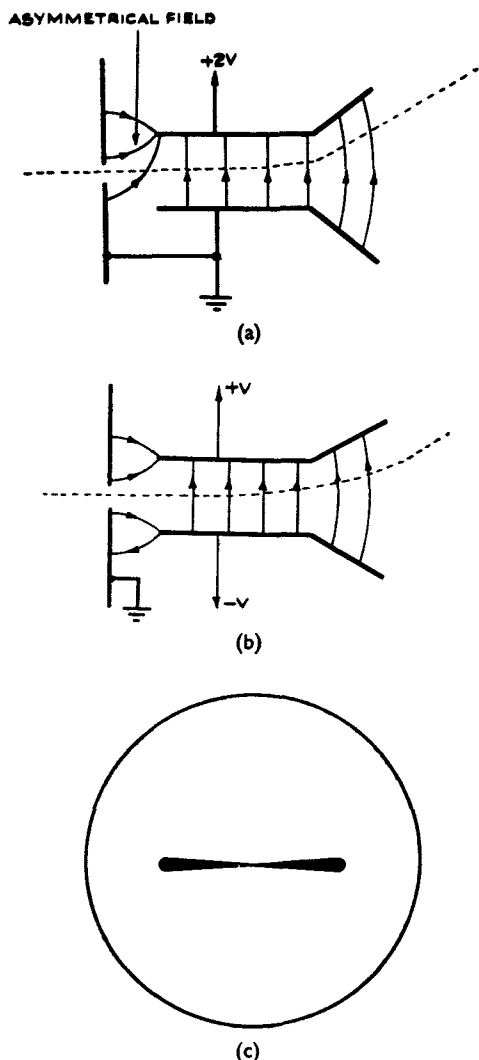


Fig. 142.—Deflection defocus

and blurred at the ends of the trace, giving the effect shown in fig. 142 (c). The origin of the term “deflection defocus” is obvious from the figure.

259. If symmetrical deflecting voltages are applied to the plates, i.e. if, when one is returned to a potential $+V$ the other is returned to $-V$, there will still be a deflecting field due to $2V$ between the plates and the beam will be deflected by the same amount as before; but the mean potential of the plates will be that of the third anode and there will be no field acting on the beam in the space

between the third anode and the plates, as shown in fig. 142 (b). Deflection defocus will thus not occur.

Trapezium distortion

260. Another form of distortion occurs when asymmetrical deflection is used. Since the mean potential of the plates varies with the signal, the axial velocity of the beam will also be variable after passing between the deflector plates, being greater when a positive signal is applied, and less when a negative signal is applied, than the velocity with which it leaves the third anode. The sensitivity of the second pair of plates is inversely proportional to the velocity of the beam when it passes between them, as explained in the section on deflection (fig. 143). We shall assume that the plates nearer the final anode are the Y plates. If a

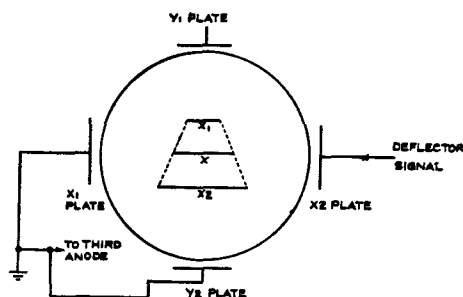


Fig. 143.—Trapezium distortion

given signal applied to the X plates produces a trace of length x when no signal is applied to the Y2 plate, a positive voltage applied to the Y2 plate will deflect the trace upwards and also increase the velocity of the beam before it reaches the X plates. The length of the trace will therefore be diminished to, say, x_1 . Similarly, if a negative voltage is applied to Y2 the trace will be shifted downwards and its length increased to, say, x_2 . The envelope of the trace as Y2 varies will therefore be a trapezium, which gives rise to the term “trapezium distortion”, and it can be seen that the picture of any waveform applied to the X plates will not be a true representation of the waveform.

261. Trapezium distortion can be overcome either by using symmetrical deflection for the X plates, or by inserting between them and the Y plates a plate with its plane perpendicular to the beam and with its potential that of the third anode. Such a plate is usually connected

internally to the third anode. The velocity of the beam is thus rendered constant after leaving this plate and the deflection produced by a given signal on the Y plates is independent of any signal on the X plates.

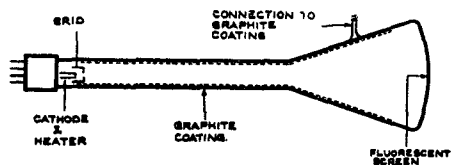


Fig. 144.—Electromagnetic CRT

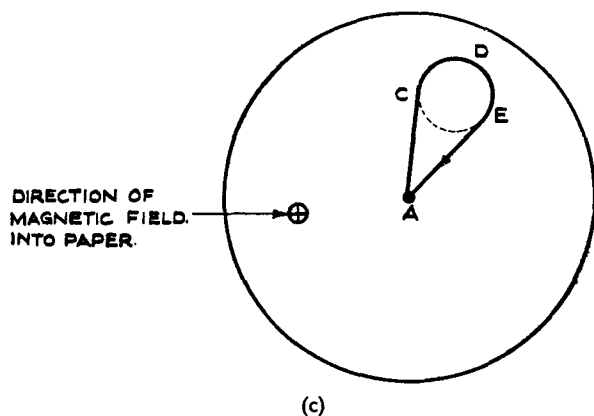
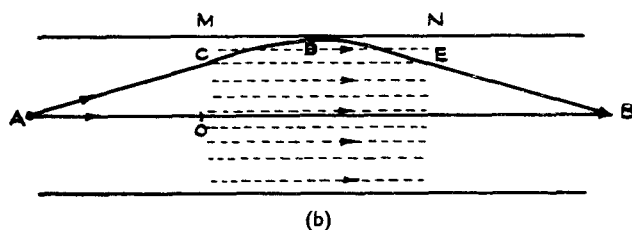
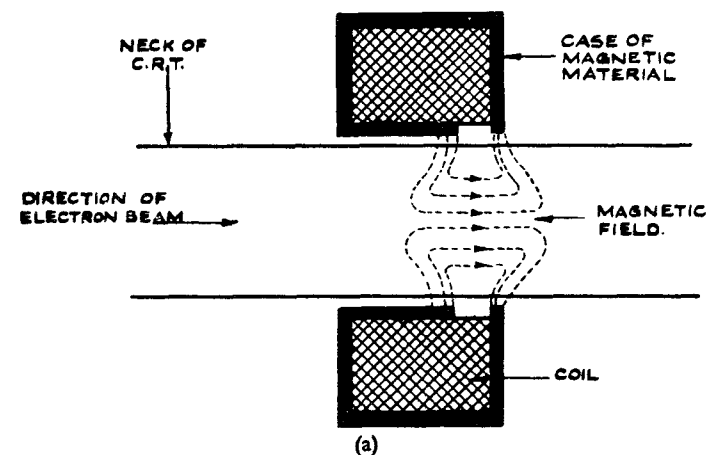


Fig. 145.—Electromagnetic focussing

Electromagnetically-focussed and deflected tubes

262. The construction of these tubes differs somewhat from that of an electrostatic tube (fig. 144). One anode only is used, this commonly being the graphite coating itself. The neck of the tube is much narrower as the focussing and deflector coils (which take the place of the focussing anode and deflector plates of an electrostatic tube) are mounted outside the neck. In an electrostatic tube the size of the beam is restricted by the close spacing of the deflector plates (about $\frac{1}{2}$ cm. apart). Thus the permissible cross-sectional area of the beam itself in the neck of an electromagnetic tube is greater than in an electrostatic tube; consequently, the light-spot can be much brighter. This fact, and the simplicity of construction, constitute the main advantages of electromagnetic over the electrostatic CRTs. The disadvantages will be mentioned later under focussing and deflection.

Focussing (fig. 145)

263. Focussing may be accomplished by means of a coil of wire, wound coaxial with the neck of the tube, through which DC is passed; this produces a magnetic field through which the beam passes. The coil is encased in magnetic material which shapes the field so that in the main it is longitudinal (i.e. is parallel with the axis of the neck). The strength of the field is proportional to the current flowing in the coil.

264. *Focussing action.*—A moving electron constitutes an electric current whose conventional direction is opposite to the direction in which the electron is moving. Thus, an electron moving in a magnetic field which is at right-angles to its direction of motion will be deflected in a direction at right-angles both to the direction of the field and the direction of the conventional current according to Fleming's left-hand rule.

265. The field in fig. 145 (a) may be idealised to that in fig. 145 (b), where it consists of a uniform longitudinal field which exists only between M and N. Consider a beam of

electrons diverging in a cone from a point A and travelling from left to right. Those electrons moving along the axis of the field will be unaffected and will describe the straight path AOB. An electron travelling along a path such as AC will meet the magnetic field at C and its velocity will have two components—one longitudinal which will carry it through the field and will be unaffected by the field, and one radial from the axis, i.e. at right-angles to the field. Viewing the system from A, along A-B (fig. 145c) this radial component of velocity causes the electron to follow the circular path CDE which is stretched out by the longitudinal component into part of a helix or spiral. The radius of the circle is dependent on the strength of the magnetic field; this is adjusted by altering the current in the focussing coil so that when the electron leaves the field at E it is travelling back towards the axis. It will follow the straight line EB, and meet the axis at B. It can be shown that for the same magnetic field any electron, no matter what the angle at which it originally diverged from the axis, will also be brought to B, i.e. the beam will be focussed on B, which is arranged to lie on the screen.

266. In the practical case of fig. 145 (a) the focus is controlled by controlling the current in the coil and by altering the position of the coil along the neck of the tube. When the coil is near the first anode, focussing is best when the spot is near the centre of the screen, and when the coil is moved towards the screen the focus is better towards the edge of the screen.

267. The disadvantage of this method of focussing is the power dissipated in the resistance of the coil by passing the focussing current through it, and the necessity for providing a transformer, rectifier, and smoothing circuits to supply the current, all of which add weight and cost to the equipment.

268. *Deflection* (see fig. 146).—Suppose a pair of coils AA' is placed on the neck of the CRT, with its axis perpendicular to the axis of the tube, between the focus coil and the screen. A current passed through the coil will produce a magnetic field across the neck of the tube, as shown in fig. 146(b). If b is the beam of electrons (moving into the paper) then by Fleming's left-hand rule it will be deflected to the left, by an amount proportional to the intensity of the field, i.e. proportional to the exciting current. If the direction of the

current is reversed then the direction of deflection will be reversed. Thus if any *current* waveform is passed through the coils the deflection of the spot will follow this waveform. Compare the effect with the application of a *voltage* waveform to a pair of deflector plates of an electrostatic CRT.

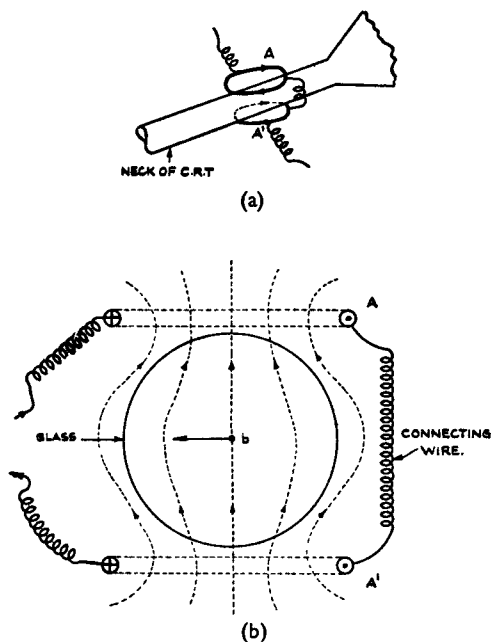


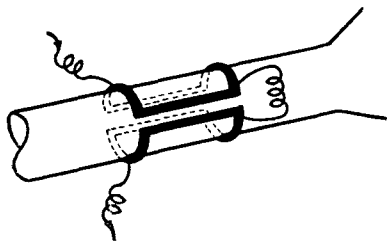
Fig. 146.—Electromagnetic deflecting system

269. In order to obtain a more uniform magnetic field across the neck of the CRT and hence a more linear relationship between deflection and current, the coils are usually wound on a rectangular former and then bent round the neck as shown in fig. 147. If deflection in two directions at right-angles is required, then another pair of coils similar to the first is placed over the first pair so as to produce a magnetic field in a direction at right-angles to that of the first.

270. The disadvantages of this system of deflection are:—

(i) It is usual to deal with *voltage* waveforms in valve circuits, and it is difficult to convert these to *current* waveforms due to the inductance of the deflecting coils.

(ii) As for the focussing coil, considerable power is dissipated in the resistance of the coils.



(a)

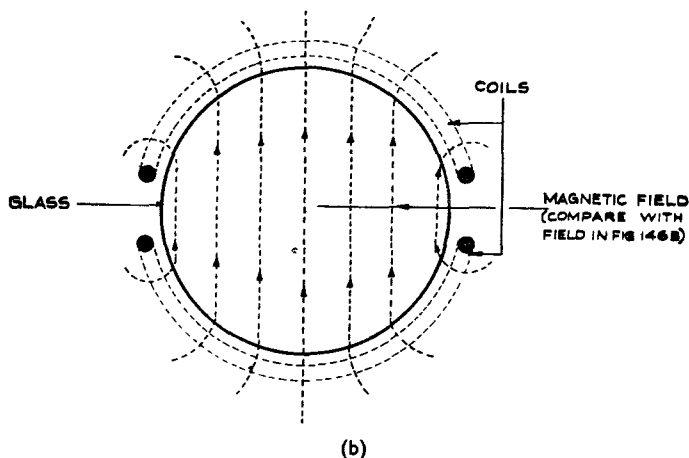


Fig. 147.—Deflector coils

Electrostatic focussing and electro-magnetic deflection

271. This system uses a tube which resembles closely an electrostatic tube, but has a deflector coil assembly mounted round the neck between the focussing anodes and the screen instead of deflector plates. The advantage is that it saves the power dissipated in the normal focussing-coil, while still allowing an intense light-spot as the size of the beam is not restricted by deflector plates mounted inside the tube. It suffers, of course, from the two disadvantages mentioned above for electro-magnetic deflection.

TIME BASE CIRCUITS FOR ELECTRO-STATIC DEFLECTION

272. The fundamental time base circuit is shown in fig. 148. The ideal time base should have a charge device which passes a constant current into the capacitor, the p.d. across which therefore rises absolutely linearly. At the end of the charging period the discharge device should discharge the capacitor instantaneously.

273. There are two possible methods of operating such an ideal time base:—

(i) Self-running, in which no external waveform need be applied to the circuit. C charges until the p.d. across it reaches a critical value, when the discharge device is automatically switched on. At the end of the discharge it switches itself off and the charging of C restarts (fig. 149).

(ii) Triggered, in which an external waveform is needed. In fig. 150, the negative-going edge of the square wave starts the charging of C while the positive-going edge starts the discharge. The circuit then remains in a stable state with C discharged until the next negative-going edge arrives.

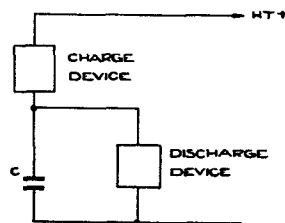


Fig. 148.—Fundamental time base circuit

274. Self-running time bases are not used in radar practice, since the start of the time base must always coincide with the firing of the transmitter. This is done either by triggering the time base from a pulse derived from the transmitter or by the use of a timing section in the equipment which controls both the transmitter firing and the time base start.

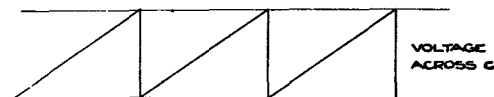


Fig. 149.—Output waveform of ideal self-running time base

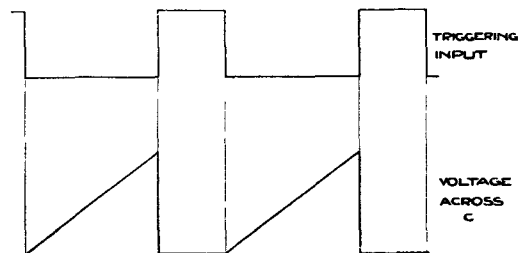


Fig. 150.—Waveforms of ideal triggered time base

275. Two simple types of triggered time base suitable for use with electro-statically deflected tubes will be considered here.

The single triode time base

276. The circuit diagram and waveforms are shown in figs. 151 and 153 respectively. An asymmetrical square wave is applied to the grid and is D.C.-restored so that the grid is at zero from D to E (fig. 153) and well beyond cut-off from B to C. Consider conditions at A. The grid is at zero and no charging current is flowing in the capacitor, the anode voltage is

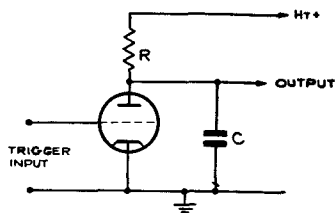


Fig. 151.—Circuit of single triode time base

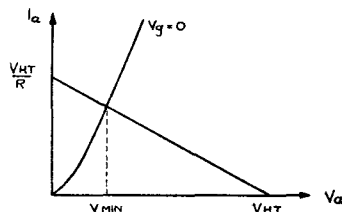


Fig. 152.—Graphical method of finding V_{min} .

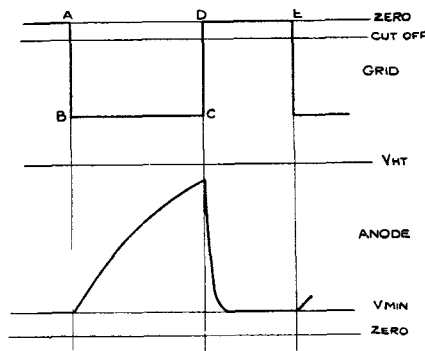


Fig. 153.—Waveforms of circuit of fig. 151

then given by the intersection of the load line corresponding to R with the zero grid volts line in the $I_a:V_a$ characteristics in the usual way (fig. 152). Call this value of anode voltage V_{min} .

(i) *A to B.* The grid moves very rapidly negative and the valve is cut-off.

(ii) *B to C.* C charges up from V_{min} towards $HT+$ on a time constant CR .

(iii) *C to D.* The grid rises abruptly to zero, thus turning the valve on hard.

(iv) *D to E.* C now discharges through the valve, the anode returning to V_{min} .

277. Notes.—(i) The time base sweep is exponential and not linear.

(ii) The linearity is improved if only the first part (say 10%) of the exponential rise is used. This means either a drastic reduction in the amplitude of the sweep or else a high value of HT voltage. If a 300-volt supply is used the amplitude must be limited to some 30 volts, while if a 200-volt sweep is required, the HT would need to be 2,000 volts.

(iii) To get a low value of V_{min} a valve of low R_a is required. This helps the discharge process since such a valve passes a heavy current when its grid is suddenly brought up to zero and C is therefore discharged quickly.

(iv) The duration of the time base sweep is fixed by the duration of BC in the input waveform.

(v) The flyback time is determined by the rate of discharge of C through the valve and DE must be long enough for this to be completed before the next sweep starts.

(vi) No use is made of the amplifying properties of the valve, which is merely an electronic switch operated by the grid waveform.

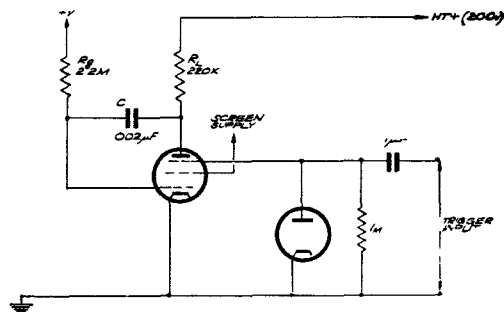


Fig. 154.—Basic circuit of suppressor triggered Miller time base

The Miller time base (suppressor-triggered)

278. The basic circuit is shown in fig. 154.

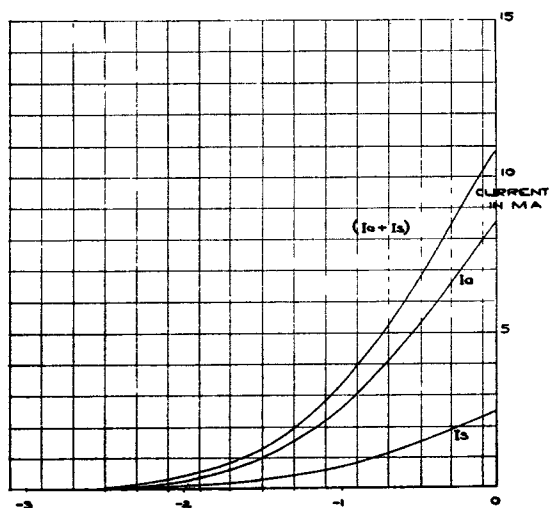


Fig. 155.—Mutual characteristic of VR.91 for screen volts = 100

It should be noted that the anode and grid are coupled by the capacitance C and that the load in the anode circuit is considerably higher than is normally used with a pentode. The amplifying properties of the valve are used to give a time base sweep which is linear to a high degree of approximation. In order to explain the operation, numerical values will be used and it is assumed that the valve is a VR91 operated at a screen potential of 100 volts.

279. The mutual characteristics with this value of screen voltage are shown in fig. 155, it being assumed for the sake of a simplified treatment that these are independent of the anode voltage so long as the anode is not bottomed.

280. The triggering input is an asymmetrical square wave, D.C. restoration being used so that the suppressor voltage does not rise above zero; the resultant waveforms are shown in fig. 156.

(i) *At A* (first stable condition).—The suppressor is held at some -100 volts, i.e. the suppressor is sufficiently negative to prevent any flow of current to the anode. The grid is held at zero by grid current and the total space current (11 mA under these conditions) flows to the screen. The anode is at H.T.+ and the capacitor C thus charged to a p.d. of 200 volts.

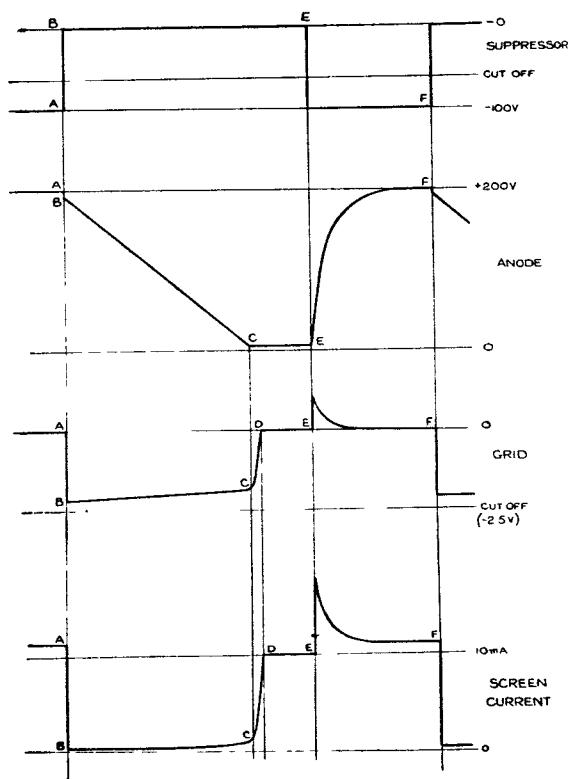


Fig. 156.—Waveforms of the suppressor triggered Miller time base

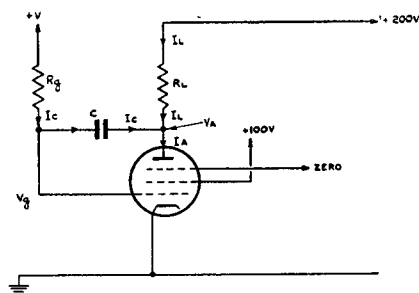


Fig. 157.—Circuit conditions of Miller time base during run-down

(ii) *A to B* (initial jump).—The suppressor is suddenly raised to zero and anode current starts to flow, the anode potential falling. This fall in anode potential is applied via C to the grid and thus the tendency of the anode current to rise is offset by a corresponding fall in grid potential. With the specified screen voltage the grid base is 2.5 volts and it is clear that the drop in anode voltage cannot exceed this value. The anode current must therefore be very small and the valve nearly cut-off, a more detailed calculation giving -2.2 volts for the grid potential at B .

(iii) *B to C* (run down).—The grid having been driven down nearly to cut-off, grid current has ceased and the current through R_g no longer flows through the grid-cathode path of the valve but flows instead into the capacitor C which discharges. During this discharge the grid rises and the anode falls, the change in voltage across C thus being shared between anode and grid. Now although the valve is nearly at cut-off the high anode load gives a gain of about 300 between grid and anode so that as the grid rises the anode falls 300 times as fast. The rise of grid voltage is therefore very small and to a first approximation the grid may be considered to be at a constant voltage. It may be shown that the anode voltage falls linearly.

281. Let V_a = instantaneous anode voltage
 V_g = instantaneous grid voltage
 then $V_a - V_g$ = instantaneous capacitor p.d.
 Referring to fig. 157 we have:—

$$I_C = \frac{V - V_g}{R_g} = \text{constant} \quad \dots (i)$$

Also

$$\begin{aligned} I_C &= C \times \text{rate of change of capacitor p.d.} \\ &= C \times \text{rate of change of } (V_a - V_g) \\ &= C \times \text{rate of change of } V_a \dots (ii) \end{aligned}$$

since V_g is regarded as constant.

Now, since I_C is constant (equation (i)), the rate of change of V_a is constant, i.e. V_a falls linearly.

Since V_g is of the order of -2 volts and V is, say, $+200$ volts, only 1% error will be made by neglecting V_g in equation (i).

Combining (i) and (ii):—

$$\text{rate of change of } V_a = \frac{V}{CR_g} \quad \dots (iii)$$

282. The anode voltage decreases linearly until the point C is reached. Here the anode bottoms and the amplifying action of the valve ceases. With the high anode load used (220K) the anode bottoms at the low potential of about 5 volts. It is interesting to examine the circuit conditions at this point. The load current (I_L in fig. 157) is 0.89 mA and the capacitor current I_C 0.09 mA, the anode current thus being 0.98 mA. To enable the valve to pass this current the grid must have risen to -1.5 volts, a change of $+0.7$ volt compared with the 195 volts drop at the anode.

(i) *C to D*.—When the valve bottoms, the anode is held at the bottoming voltage and the capacitor current I_C then enables the grid to rise quickly until it is caught by grid current.

(ii) *D to E* (second stable state).—The grid is held at zero and the anode bottomed. The space current is now 11 mA as at A , of which 0.89 mA flows to the anode and the remainder to the screen.

(iii) *E to F*.—At E the suppressor returns to -100 volts, cutting off the anode current. C now recharges through R_L and the grid-cathode path of the valve, the anode rising to $HT+$ on a time constant CR_L and the usual pip appearing on the grid waveform.

283. Notes

(i) Time of run-down

It has been shown that:—

$$\text{rate of change of } V_a = \frac{V}{CR_g} \quad \dots (iii)$$

If ΔV_a is the total change of anode volts during run-down and T is the time taken, the rate of change of V_a is $\Delta V_a/T$, so that from (iii)

$$T = \frac{\Delta V_a}{V} \cdot CR_g \quad \dots (iv)$$

Now ΔV_a is equal to the HT voltage (V_{HT}) less the initial drop of about 2 volts and the bottoming voltage (5 volts). So that approximately $\Delta V_a = V_{HT}$ and

$$T = \frac{V_{HT}}{V} \cdot CR_g \quad \dots (v)$$

(ii) Time of flyback

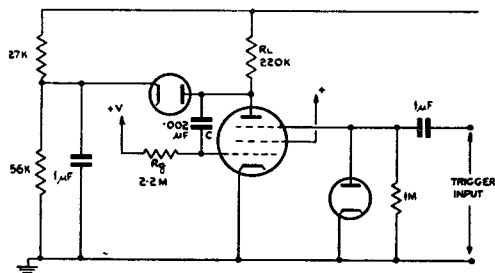
The anode returns to $HT+$ on a time constant CR_L and has risen to within 1% of the $HT+$ level in a time equal to $5 CR_L$. This is taken as the time of flyback.

(iii) Screen current waveform

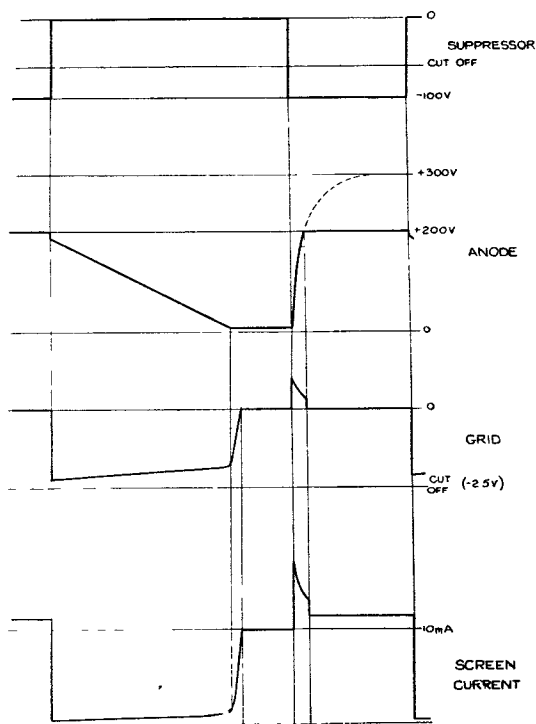
With the suppressor at -100 volts, at A the full space current of 11 mA flows to the screen. During the run-down the suppressor is at zero and the screen current will therefore be about $1/3$ of the anode current, rising sharply at the end of the run-down as the anode bottoms and the grid potential comes up to zero.

284. It has been shown that from D to E the screen current is about 10 mA. During the flyback the grid waveform shows a positive pip and therefore a corresponding positive pip appears on the screen current waveform.

If now a resistor is placed in the screen lead the resultant screen voltage waveform will be the same as that of the screen current but inverted and it will be seen that this will have a nearly flat-topped positive-going portion during the run-down. If this is applied to the grid of the cathode ray tube the beam will be turned on during run-down (i.e. during the time base sweep) and will be blacked out during the remainder of the cycle. Some squaring of the screen waveform is usually needed before using it in this manner so that there is no brightness variation during the sweep.



(a)



(b)

Fig. 158.—Circuit and waveforms of suppressor triggered Miller time base

Use of anode-catching diode

285. In order to get good linearity of run-down and good bottoming of the anode it is essential to keep the anode load high, and this may unduly prolong the flyback time.

286. A considerable reduction in this time may be obtained by the use of an anode-catching diode (fig. 158). The HT supply is increased to 300 volts and the cathode of the diode is taken to about $2/3$ HT. Considering the waveforms of fig. 156 it will be seen that at A, when the suppressor is at -100 volts, the diode is conducting, the current through it and the anode load being 0.45 mA. The run-down now starts from $+200$ volts, the initial drop at the grid being to -1.8 volts. At the end of the run-down the drop across the load is 295 volts, the anode current 1.43 mA and the grid voltage -1.35 volts.

287. During the flyback the anode is rising to $+300$ volts on the time constant CR_L , but it can rise only to $+200$ volts before the diode conducts and prevents any further rise. In one time constant (i.e. CR_L) the anode has risen to 63% of the HT voltage, i.e. to 189 volts, and will therefore reach 200 volts very shortly afterwards. In fact, for all practical purposes, the flyback time may be taken as CR_L and the use of the diode has therefore reduced this time by a factor of 5.

288. Since the charging of C stops as soon as the anode is caught by the diode, grid current ceases and the grid returns rapidly to zero.

Variation of run-down time and amplitude

289. From equation (v) the run-down time T is given by:—

$$T = \frac{V_{HT}}{V} \cdot CR_g$$

This was calculated on the assumption that the run-down started from V_{HT} . When an anode-catching diode is used, the rundown starts from the voltage to which the cathode of the diode is taken, say, V' . The run-down time is then

$$T = \frac{V'}{V} \cdot CR_g \dots\dots\dots (vi)$$

Clearly, any one of the four quantities on the right-hand side may be changed to vary the time of run-down.

290. The amplitude of the sweep is of course the total change of anode volts during the run-down, and this is most easily varied by taking the cathode of the anode-catching diode not to a fixed potential but to a potentiometer across the HT rails. This will also vary the

time of flyback, but if a fixed resistor is included between the top of the potentiometer and HT+ so that the cathode of the diode can never be taken to a voltage greater than $\frac{2}{3}$ HT the flyback time will never exceed CR_L .

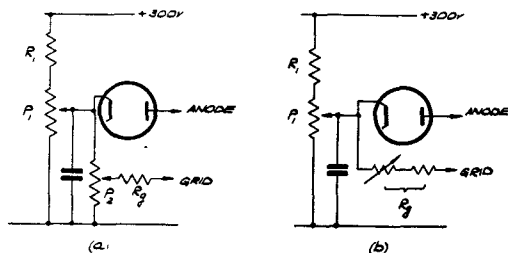


Fig. 159.—Two methods of controlling amplitude duration of run-down

291. Referring to equation (vi), if the run-down time is to be independent of amplitude the ratio $\frac{V'}{V}$ must remain constant as V' is varied. Two possible circuit arrangements are given in fig. 159. In (a) R_g is returned to the potentiometer P_2 , the setting of which varies the ratio $\frac{V'}{V}$ and hence the run-down time. Since this ratio depends only on the setting of P_2 , the amplitude control P_1 may be varied without affecting the duration of the run-down. In fig. 159 (b) the ratio $\frac{V'}{V}$ is made equal to unity by returning R_g to the cathode of the anode-catching diode. The run-down time is then varied by making R_g variable.

292. *Example.*—It is required to design a Miller time base to fulfil the following requirements:—

- (i) Amplitude variable from 100 to 200 volts.
- (ii) Duration of run-down 500 to 1000 microseconds.
- (iii) Flyback time not to exceed 50 microseconds.

For good bottoming R_L is taken as 220K. Then CR_L must not exceed 50 μ secs.

$$\therefore C < \frac{50 \times 10^{-6}}{220 \times 10^3} = 227 \text{ pF.}$$

A suitable value would be 200 pF.

Choosing the circuit of fig. 159 (a) the minimum run-down time is obtained for $\frac{V'}{V} = 1$, so that

$$500 \times 10^{-3} = CR_g$$

which gives $R_g = 2.5 \text{ M.}$

For the maximum run-down time $\frac{V'}{V} = 2$ or $V = \frac{1}{2} V'$ and the voltage to which R_g is returned must therefore lie between V' and $\frac{1}{2} V'$. The variation of the potential of the anode-catching diode (amplitude control) is similarly restricted. The complete circuit is shown in fig. 160.

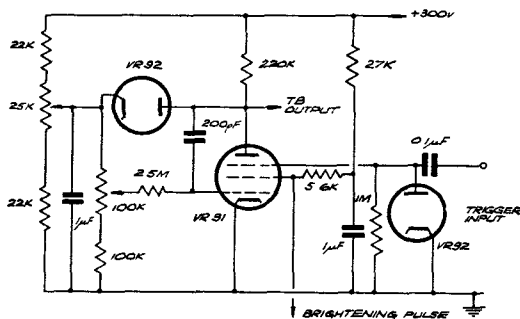


Fig. 160.—Complete circuit diagram of Miller time base

The output impedance of the Miller valve

293. One very important advantage of the Miller time base circuit is its low output impedance, a property which is extremely useful when it is desired to feed fast time base sweeps to the CRT plates.

294. To show that the stage has a low output impedance, proceed as follows:—

Suppose that during the run-down a small voltage change ΔV is applied to the anode from some external source. Since grid and anode are connected by the capacitor C , this voltage change is also applied to the grid, causing a change in anode current equal to $g_m \Delta V$. This incremental anode current must be drawn from the source which causes it, so that the source must "look into" a resistance of $\frac{\Delta V}{g_m \Delta V}$ or $\frac{1}{g_m}$. But the impedance into which the source looks is by definition the output impedance of the stage,

and this output impedance is therefore resistive and of value $\frac{1}{g_m}$, the same as that of a cathode follower.

295. In practice, the output impedance of a VR91 used in a Miller time base is somewhat higher than a similar valve used in the conventional cathode follower circuit, since the Miller stage works at a lower anode current and therefore in a region where the mutual conductance is low. Even so, the output impedance is not more than some 500 ohms.

BALANCED OUTPUT STAGES

296. It has been seen in para. 258 that if the mean potential of the deflector plates of an electrostatically-deflected CRT varies, then the electrostatic fields used to focus the electron beam are disturbed, and deflection defocussing occurs. Such variation of the mean potential is produced when the deflection of the beam is obtained by applying the deflecting voltages to only one plate of a pair. The remedy is to apply waveforms of equal amplitude, but opposite polarity to both plates. Such waveforms are known as paraphase or balanced waveforms. Their use has the further advantage of doubling the deflection of the beam.

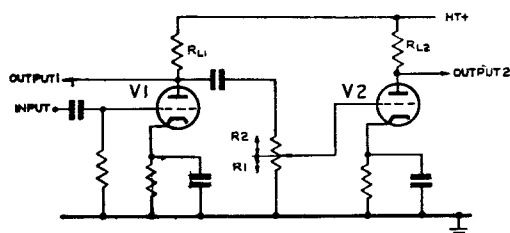


Fig. 161.—Paraphase circuit

297. Circuits producing paraphase waveforms may be termed balanced output stages; the paraphase circuits and the long-tailed pair are representative.

298. Fig. 161 shows the paraphase circuit. V1 is a normal resistance loaded amplifier, the output of which feeds one deflector plate. V2 provides the phase-inverted output to the other plate, and to allow for the amplification of V2, its grid is fed from V1 via the voltage divider R1, R2. Equality of output is obtained by making

$$\frac{R1}{R1 + R2} \cdot A2 = 1, \text{ where } A2 \text{ is the numerical gain of } V2.$$

V1 and V2 are shown as triodes for simplicity, but will usually be pentodes because of their smaller Miller effect. If the circuit is used to provide paraphase time base voltages, V1 will usually form the time base generator.

299. Disadvantages of the simple paraphase circuit are as follows:—

(i) The operation of the circuit depends on the value of A2, which varies with the specimen and the age of the valve V2.

(ii) Distortion is produced by non-linearity of the valve characteristics. This may be reduced by omitting the by-pass capacitors in the cathode circuits, i.e. by introducing current negative feedback.

(iii) Capacitance in shunt with the outputs can cause distortion of the waveforms (para. 162).

These disadvantages are overcome to a large extent by the use of the floating paraphase circuit.

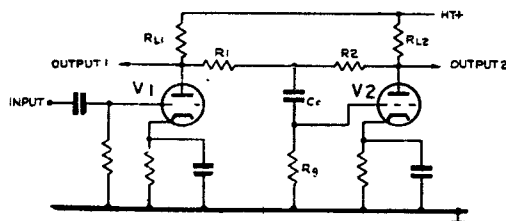


Fig. 162.—Floating paraphase circuit

Floating paraphase

300. Fig. 162 shows one form of the floating paraphase circuit in which triodes are again shown for simplicity. The output of V1 is fed to the grid of V2 via R1, Cc and Rg, and the phase-inverted output at the anode of V2 is fed back to the grid of V1 via R2, Cc, and Rg. The input to V2 is thus derived from both outputs, and by proper choice of R1 and R2, equal outputs may be obtained.

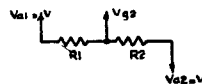


Fig. 163.—Voltage dividers of floating paraphase

301. Conditions for balanced outputs are determined thus:—It is assumed that Rg is much larger than R1 and R2.

(i) If Va1 rises by V, and Va2 falls by the same amount, as in fig. 163, then relative to Va1, Va2 falls by 2V. The potential divider

R_1, R_2 causes V_{g2} to fall by $\frac{R_1}{R_1 + R_2} 2V$ with respect to V_{a1} . Since V_{a1} rises by V , the actual V_{g2} rise is

$$V - \frac{R_1}{R_2 + R_1} 2V \\ = \frac{R_2 - R_1}{R_2 + R_1} V$$

(ii) If the numerical gain of V_2 is A_2 , then

$$A_2 \frac{R_2 - R_1}{R_2 + R_1} V = V$$

$$\therefore \frac{R_2 - R_1}{R_2 + R_1} = \frac{1}{A_2}$$

$$\therefore \frac{R_1}{R_2} = \frac{A_2 - 1}{A_2 + 1} \\ = 1 - \frac{2}{A_2 + 1}$$

(iii) From this expression it can be seen that if $A_2 \gg 1$,

$$R_1 \simeq R_2$$

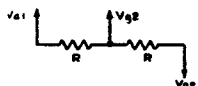


Fig. 164.—Voltage dividers with equal arms

Inequality of outputs if $R_1 = R_2 = R$

302. If the outputs are as shown in fig. 164, then the change of V_{g2} is the mean of the output changes,

that is V_{g2} rises by $\frac{1}{2} (V_{a1} - V_{a2})$.

As before,

$$A_2 \cdot \frac{1}{2} (V_{a1} - V_{a2}) = V_{a2}$$

therefore

$$\frac{V_{a1}}{V_{a2}} = 1 + \frac{2}{A_2}$$

303. Example.—

For a VR91 pentode, $g_m = 5$ mA/volt. If the anode load of V_2 is 33K, then $A_2 = g_m R_L$
 $= (5 \times 10^{-3}) \times (33 \times 10^3)$

i.e. $A_2 = 165$.

$$\therefore \frac{V_{a1}}{V_{a2}} = 1 + \frac{2}{165} \simeq 1.01$$

Thus if A_2 is large, and $R_1 = R_2$, then the ratio of outputs is approximately unity, and is almost independent of A_2 , i.e. of V_2 and R_{L2} . Consequently it is customary to make R_1 and R_2 equal.

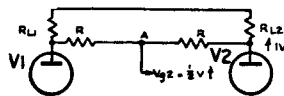


Fig. 165.—Output impedance of floating paraphase

Output impedance of V_2

304. The output impedance may be derived by calculating the current drawn from a source of voltage which raises the anode voltage of V_2 (in this case by, say, 1 volt).

We shall assume that

$$R_1 = R_2 (= R, \text{ say}),$$

$$R \gg R_{L1}$$

and $R \ll R_g$

e.g. $R = 220K, R_{L1} = 22K, R_g = 2.2M$.

305. From fig. 165 it may seem that when V_{a2} is instantaneously raised by 1 volt, then the voltage at point A rises by approximately $\frac{1}{2}$ -volt. This sudden rise is transferred by the $C_0 R_g$ coupling to the grid of V_2 . If V_2 is a pentode, then the anode current of V_2 rises by $\frac{1}{2} \cdot g_m$ where g_m is the mutual conductance of V_2 . The output impedance, Z_o , is the ratio of the applied voltage change to the current change produced.

$$\therefore Z_o = \frac{1 \text{ volt}}{\frac{1}{2} \cdot g_m \text{ amps.}}$$

$$\therefore Z_o = \frac{2}{g_m} \text{ ohms.}$$

Example—

For the case of the VR91 considered in the last example

$$Z_o = \frac{2}{5 \times 10^{-3}} \text{ ohms}$$

$$\therefore Z_o = 400 \text{ ohms.}$$

306. This low output impedance reduces the effect of stray capacitance on the output waveform of V_2 . If V_1 consists of a normal pentode amplifier, it will normally have a fairly high output impedance; however, should V_1 form a Miller time base, its output impedance will be low ($\frac{1}{g_m}$) and neither output will be appreciably affected by strays.

307. The voltage negative feed-back from anode to grid of V2 makes the output of V2 very linear. Capacitors (50–100 pF) may be used in place of the resistors R1 and R2 as the cross-coupling components.

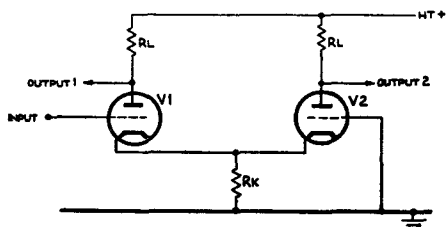


Fig. 166.—Basic circuit of long-tailed pair

Long-tailed pair

308. The basic circuit of a long-tailed pair is shown in fig. 166. Triodes or pentodes may be used; triodes are shown in the diagram, but it is assumed for simplicity of analysis that pentodes are employed.

309. To understand the general operation, suppose that the input causes V_{g1} to rise; I_{a1} then rises, causing V_{a1} to fall. Also, the rise of I_{a1} through R_K raises V_K . Since the grid of V2 is tied to earth, this rise of V_K is equivalent to a fall of V_{gk2} , which makes I_{a2} fall and V_{a2} rise. Thus V_{a1} and V_{a2} move in opposite directions. The action is, however, complicated by the fact that the rise of I_{a1} is accompanied by a fall of I_{a2} , both currents affecting the voltage V_K .

310. The ratio of the outputs can be obtained in the following way. Suppose that the input to V1 changes so that V_K rises by 1 volt (fig. 167). This is equivalent to a 1-volt fall of V_{gk2} , which causes I_{a2} to fall by g_{m2} , where g_{m2} is the mutual conductance of V2. But the increase of total cathode current ($I_{a1} + I_{a2}$) through R_K required to produce a 1-volt rise of V_K is $\frac{1}{R_K}$. Hence I_{a1} must have increased by $\left(\frac{1}{R_K} + g_{m2}\right)$. The ratio of the current changes is therefore

$$\frac{I_{a2}}{I_{a1}} = \frac{g_{m2}}{\frac{1}{R_K} + g_{m2}} = \frac{1}{1 + \frac{1}{g_{m2} R_K}}$$

311. This ratio approximates to unity if g_{m2} and R_K are large. Increase of R_K reduces the effective H.T., causing the maximum undis-

torted outputs to fall; a compromise must therefore be made. The ratio of $\frac{I_{a2}}{I_{a1}}$ does not depend on the characteristics of V1.

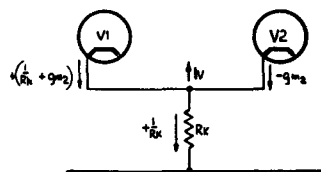


Fig. 167.—Current changes in long-tailed pair

312. The value of R_K chosen from the above considerations will usually provide too large a bias if the circuit is arranged as in fig. 166. This may be avoided by returning the grids to suitable positive potentials derived from resistance chains R1, R2 and R3, R4 across the H.T. supply (fig. 168).

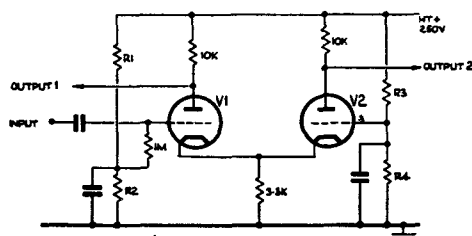


Fig. 168.—Typical long-tailed pair

Balanced shift control

313. The long-tailed pair can be provided with a balanced shift control by varying the grid potential of V2, balanced movements of V_{a1} and V_{a2} being obtained because an input to V2 produces a similar effect as an input to V1. The outputs must be D.C. connected to the deflector plates and the normal shift circuits omitted.

LINEAR TIME BASE CIRCUITS FOR ELECTROMAGNETICALLY DEFLECTED TUBES

314. In order to produce a linear time base electromagnetically, the magnetic deflecting field must vary linearly with time, and the deflecting coils must therefore carry a current varying in a similar manner. The coils may be approximately represented by inductance L in series with a resistance R_L and the equivalent circuit of a valve with the coils as anode load may then be taken as in fig. 169.

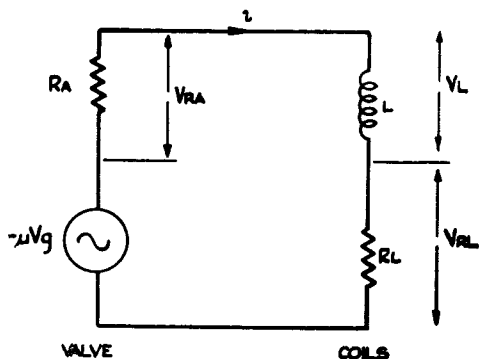


Fig. 169.—Equivalent circuit of stage-feeding coils

315. Consider that a uniformly rising current i flows in the circuit (top waveform of fig. 170). The voltage developed across R_a by the flow of current i through it equals iR_a , and therefore also rises uniformly. The voltage V_L equals $L \frac{di}{dt}$; during the period in which the rate of

change of current with time $\frac{di}{dt}$ is constant, V_L is therefore a constant voltage, rising instantaneously to this value at $t = 0$. V_{RL} equals iR_L and therefore rises uniformly.

The generator voltage μV_g required to produce the uniformly rising current is given by

$$-\mu V_g = V_{Ra} + V_L + V_{RL}$$

and is of the pedestal shape shown in the bottom diagram of fig. 170. The required grid input V_g will have the same shape.

316. *Pentode method*—If we use a pentode with a very large R_a ,

$$V_{Ra} \gg V_L + V_{RL}$$

$$\therefore \mu V_g \approx V_{Ra}$$

and consequently the required grid input is a linearly rising voltage. A current of 50-100 mA through the coils is necessary because of the magnetic field strengths required, so a power pentode or tetrode must be used. The R_a values of power pentodes and tetrodes are relatively low (20K), but may be effectively increased by means of current negative feedback. This type of feedback is most easily introduced by the use of an unbypassed cathode resistor. The employment of negative feedback has the additional advantage of reducing the distortion produced by the valve due to the curvature of its characteristics.

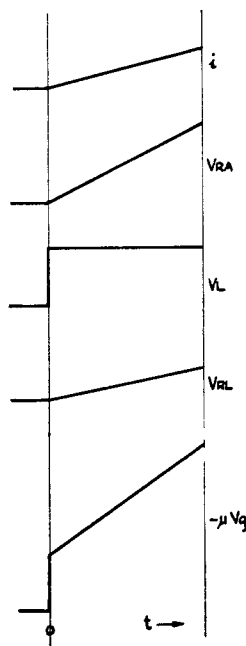


Fig. 170.—Waveforms for fig. 169

Fig. 171 shows a suitable current, the input to which may be supplied by any linear voltage time base generator.

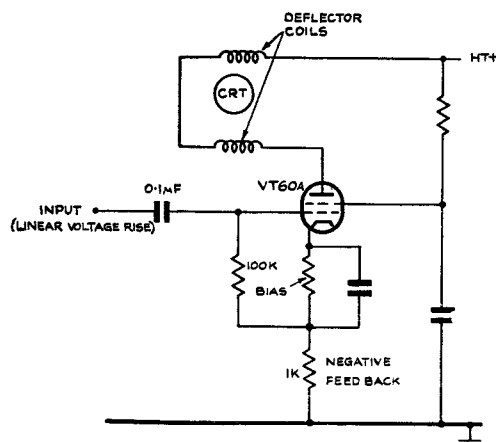


Fig. 171.—Typical tetrode output stage

317. *Pedestal voltage method*—In general, an output stage of any value of R_a may be used, provided that the correct pedestal voltage input is applied. A cathode follower type of output stage is frequently employed, because the voltage negative feedback in such a stage

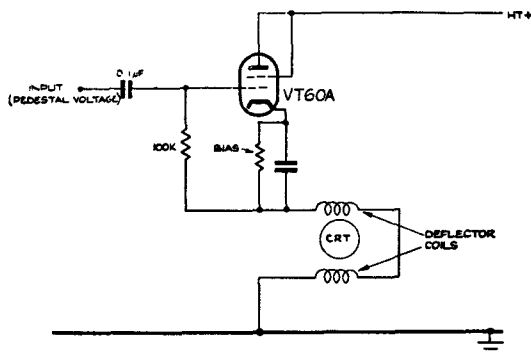


Fig. 172.—Cathode follower output stage

reduces the distortion due to the valve. Such a stage is shown in fig. 172 and its equivalent circuit in fig. 173. This equivalent circuit resembles that of fig. 170 with the exceptions that (i) the output impedance $\frac{1}{g_m}$ of the cathode follower replaces R_a , and (ii) the generator voltage is the required input V_s . The required input is therefore the sum of the voltages across the three circuit elements.

318. *Example:—*

$$g_m = 5 \text{ mA/V}$$

$$L = 100 \text{ mH}$$

$$R_L = 300 \Omega$$

Linear current rise of 50 mA in 100 μsec .

$$\begin{aligned} (V_{\max}) \frac{1}{g_m} &= i_{\max} \cdot \frac{1}{g_m} \\ &= 50 \times 10^{-3} \times \frac{1}{5 \times 10^{-3}} \\ &= 10 \text{ volts} \\ V_L &= L \frac{di}{dt} \\ &= 100 \times 10^{-3} \times \frac{50 \times 10^{-3}}{100 \times 10^{-6}} \\ &= 50 \text{ volts} \end{aligned}$$

$$\begin{aligned} (V_{\max}) R_L &= i_{\max} \cdot R_L \\ &= 50 \times 10^{-3} \times 300 \\ &= 15 \text{ volts} \end{aligned}$$

Waveforms of the above voltages together with that of their sum, i.e. the required input, are shown in fig. 174.

Pedestal waveform generator

319. If a square triggering waveform is applied to the circuit of fig. 175 then a pedestal waveform output with an approximately linear top may be obtained. The triggering waveform

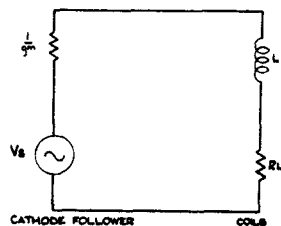


Fig. 173.—Equivalent circuit of cathode follower output stage

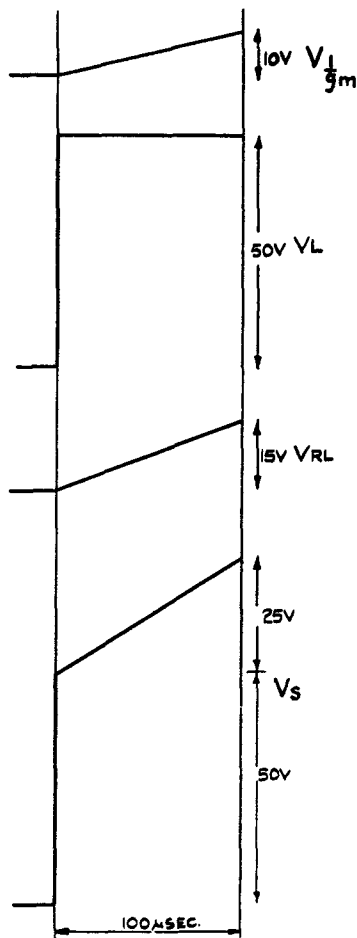


Fig. 174.—Cathode follower waveforms

should be of sufficient amplitude to cut the valve on and off; the treatment may be simplified by regarding the valve as a resistanceless switch. The stages in the operation for the particular values shown in the figure are as follows, waveforms being as shown in fig. 176.

(i) Initial conditions: Valve switch on, C uncharged.

$$\therefore V_0 = 0, V_C = 0, V_{R2} = 0.$$

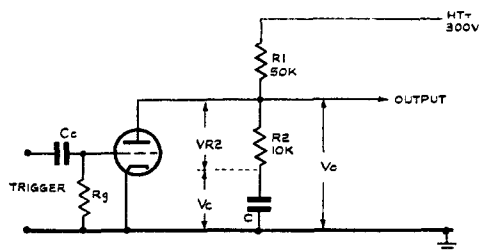


Fig. 175.—Pedestal waveform generator

(ii) Valve switched off. C cannot charge instantaneously.

$$\therefore V_{R1} = V_{R2} = 300$$

$$V_{R2} = \frac{R2}{R1 + R2} \cdot 300$$

$$= \frac{1}{6} \cdot 300$$

$$= 50V.$$

$$\therefore V_0 = V_C + V_{R2}$$

$$= 0 + 50$$

$$= 50V.$$

i.e. V_0 jumps from 0 to 50V at the instant the valve is switched off.

(iii) C charges through $R1$ and $R2$, say to 30V.

$$V_{R1} + V_{F2} = 300 - 30$$

$$= 270$$

$$\therefore V_{R2} = \frac{1}{6} \cdot 270$$

$$= 45V$$

$$\therefore V_0 = 30 + 45$$

$$= 75V.$$

(iv) Valve switched on, $V_0 = 0$. C cannot discharge instantaneously. Since $V_C + V_{R2} = V_0 = 0$, $V_{R2} = -V_C$

$$= -30V.$$

i.e. V_{R2} jumps from +45V to -30V when the valve is switched on.

(v) C discharges through $R2$ and the valve, V_C and V_{R2} returning to zero. V_0 remains at zero.

Self-capacitance of deflector coils

320. The growth of current in the deflector coils is affected by their self-capacitance. This effect becomes of importance in the case of

high speed scans for short range displays, but these further complications will not be discussed.

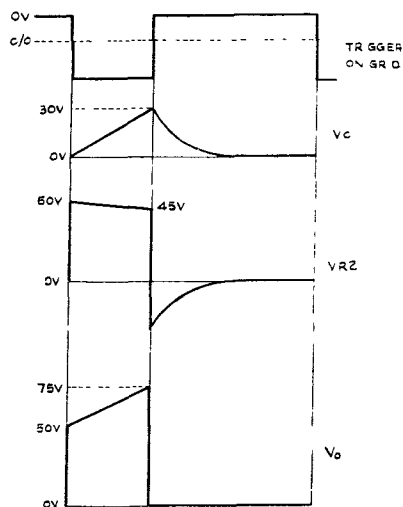


Fig. 176.—Waveforms for pedestal waveform generator

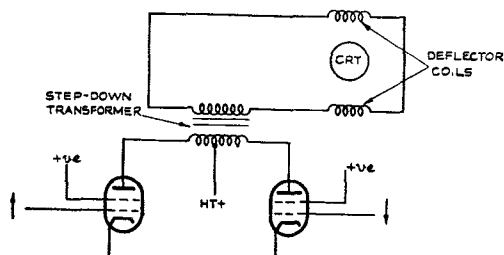


Fig. 177.—Push-pull output stage with step-down transformer

Push-pull output stage, with step-down transformer

321. It has already been stated that a heavy current varying from say 50–100 mA is required in the deflector coils, and this large current variation may cause the HT voltage to fluctuate. The difficulty may be overcome by using a push-pull output stage, as shown in fig. 177 for tetrode feed; for the valve current waveforms are then paraphase and the total current is steady. If the coils are fed from the secondary of a voltage step-down (i.e. current step-up) transformer, then the current drawn from the supply is reduced, and a slight variation of the drain has negligible effect on the voltage.

Application of shift

322. So far methods of producing a current waveform which will deflect the electron spot linearly across the screen of the CRT have been considered, but no mention has been made of the starting position of the trace. The deflection from the centre of the screen of the spot at the start of the trace is determined by the current flowing in the deflector coils at this instant. In general, this starting position will not be suitable, but a shift can be applied to the trace by passing an additional steady current through the deflector coils. The circuit used is shown in fig. 178. It can be seen that both time base and shift currents pass through the deflector coils. The permeability of the transformer core would be reduced if a DC current were to flow in the secondary; the large electrolytic capacitor C, which passes the time base current, blocks the shift current from the secondary. The large choke L passes the steady shift current but prevents the varying time base current flowing into the shift circuit.

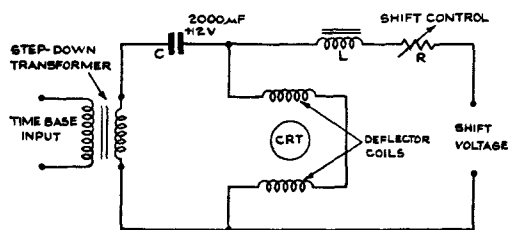


Fig. 178.—Application of shift current

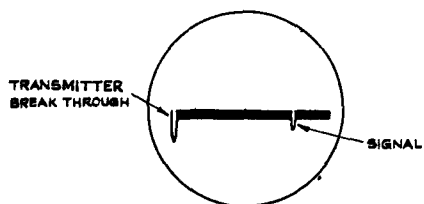


Fig. 179.—Range amplitude display

RANGE AMPLITUDE DISPLAY

323. Fig. 179 shows a range amplitude display; a time base waveform is fed to the X deflector plates to give the horizontal sweep, and the signals are fed to the Y plates to give vertical deflections. Since the time base sweep starts at the instant at which the transmitter fires, the transmitter pulse appears as a large deflection at the beginning of the trace. The separation between the transmitter break-through and an echo gives a measure of the time taken for the pulse to travel out to the target, be reflected and travel back to the

receiver. This time interval can be used to determine the range of the target. In practice, the time base is calibrated to give this range directly.

324. It is necessary to prevent signals which are received during the flyback period of the time base from being displayed; otherwise signals from targets at a distance greater than the range corresponding to the length of the forward sweep of the time base would appear to be from targets at much smaller ranges. Therefore a "bright-up" or "black-out" waveform is applied to the grid or cathode of the CRT; a square waveform of appropriate polarity (fig. 180), is used which switches on electron current during the forward stroke, and cuts off the tube during the flyback period.

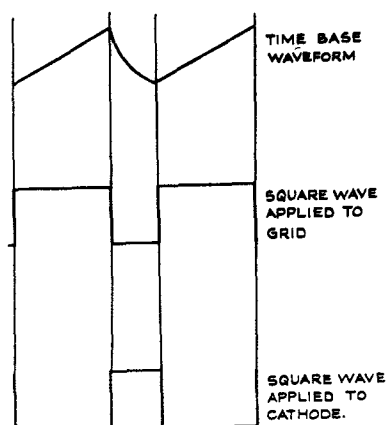


Fig. 180.—Bright-up or black-out waveforms

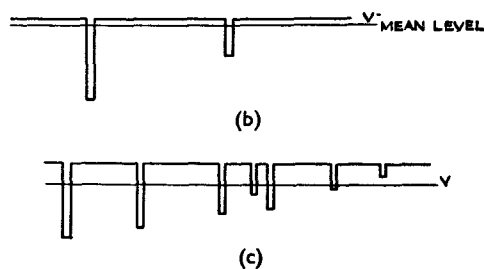
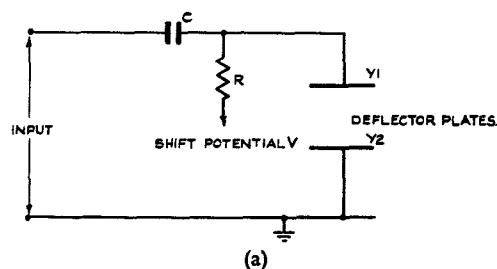


Fig. 181.—Effect of DC restoration

325. In the absence of signals the spot moves uniformly across the screen, but when a signal is applied to the deflector plates the spot moves much more rapidly, causing the signal to be less bright than the undeflected parts of the trace. This can be overcome by applying a fraction of the signal voltage to the grid of the CRT to increase the brilliance for the duration of the signal.

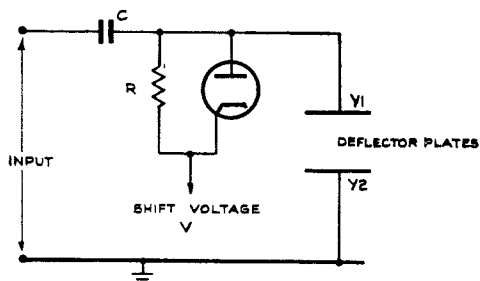


Fig. 182.—Input circuit to signal deflector plates

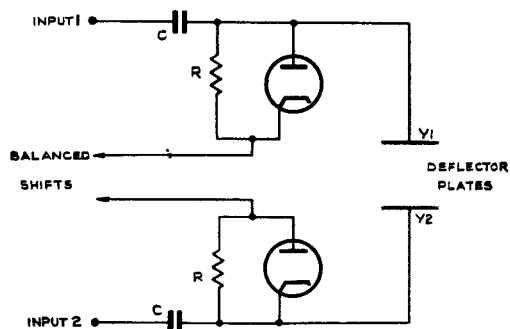


Fig. 183.—Input circuit to deflector plates for balanced signals

DC restoration

326. If the signals are applied to the Y deflector plate through a long CR coupling, as in fig. 181 (a), then the level of the trace will depend on the number and amplitude of the signals being received. For suppose that the signals shown in fig. 181 (b) are applied to the plates; then the waveform across R will have its mean level at the voltage to which R is returned, namely the shift potential. In this case the undeflected parts of the trace are at a level determined by a voltage differing little from the shift potential. Now let the signals received be increased to the number shown in fig. 181 (c). The mean level of the applied signals is lower, and consequently the un-

deflected parts of the trace are at a higher level than in the previous case. This movement of the trace can be prevented by connecting a DC restoring diode across the resistor R as shown in fig. 182. The diode restores the upper level of the signal waveform to the shift voltage, so that the base line is held at a position corresponding to the shift. A suitable circuit to use when balanced deflecting voltages are required is shown in fig. 183.

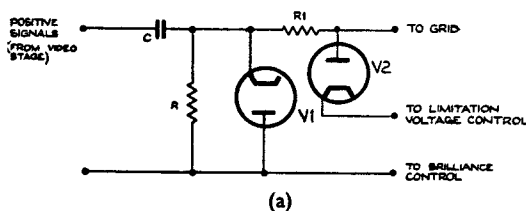


Fig. 184.—Intensity modulation

INTENSITY MODULATION OF CRT

327. In display systems using intensity modulation the presence of an echo is shown by a bright spot on the screen of the CRT. Generally the brilliance is turned down so that in the absence of signals the trace is just visible. After amplification signals received from targets are fed to the grid of the CRT to increase the brilliance. Unless suitable precautions are taken, the brightness of the spot varies with the signal strength, which depends on the range of the target. This variation of brightness makes the detection of weak signals more difficult. If the gain of the receiver is turned up to make the weak echoes brighter, then the strong signals are also amplified, resulting in brilliance defocussing of the echoes from nearby targets. This defocussing is prevented by not allowing the amplitude of the signals applied to the grid of the CRT to be greater than the optimum value.

328. The mean level of the positive signals fed to the grid of the CRT via a long CR coupling varies with the number and amplitude of the signals, and so at the grid the signals rise from a voltage level varying with the input. This level is stabilised by connecting a DC restoring diode across the resistor R. Having thus fixed the lower level of the signals at the voltage to which the grid is returned, their amplitude can be limited to the required value, say 6 volts, by connecting in the circuit a diode limiter with cathode held 6 volts positive with respect to the anode of the restoring diode. A suitable circuit for feeding signals to the grid is shown in fig. 184 (a), V1 being the restoring and V2 the limiting diode. Fig. 184 (c) shows the DC restored and limited signals applied to the grid. These signals are of equal strength with the exception of the very weak signals having insufficient amplitude to reach the limiting voltage.

329. As in the case of deflection modulation (para. 323), a black-out waveform is applied to either the grid or cathode of the CRT to cut off the tube during the flyback.

PLAN POSITION INDICATOR (PPI) DISPLAYS

330. In a PPI display the presence of a target is indicated by a bright spot on the screen. The distance from the centre of the screen to the spot is a measure of the range of the target, and the bearing is indicated by the particular radius on which the spot appears. A radial time base rotating in synchronism with the aerial system is employed, so that when energy is sent out from the transmitter in a particular direction, the time base is then sweeping along the radius corresponding to that direction. Now the energy is transmitted in a beam of finite width, and so a target is illuminated not momentarily but for the whole time that the beam is sweeping through it. Consequently the response on the CRT cannot be a point, but will take the form of an arch subtending at the centre of the screen an angle equal to the angular width of the beam.

Electrostatically deflected PPI

Production of rotating radial time base

331. The rotating radial time base is produced by applying appropriate voltage waveforms to both the X and Y deflector plates of an electrostatic CRT.

332. Take the radius, OA, drawn vertically upwards from the centre of the screen, as reference direction (fig. 185). A time base

sweep along OA is obtained by applying a positive going sawtooth voltage waveform to the Y1 deflector plate, the other plates being held at zero. Similarly a sweep of equal length along the radius OB, perpendicular to OA, is obtained by holding X2 and the Y plates at zero, and applying the same waveform to the X1 plate. For sweeps along OA' and OB' negative going sawtooth voltages must be applied to the Y1 and X1 plates respectively.

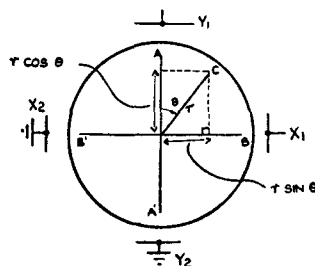


Fig. 185.—Analysis of radial scan

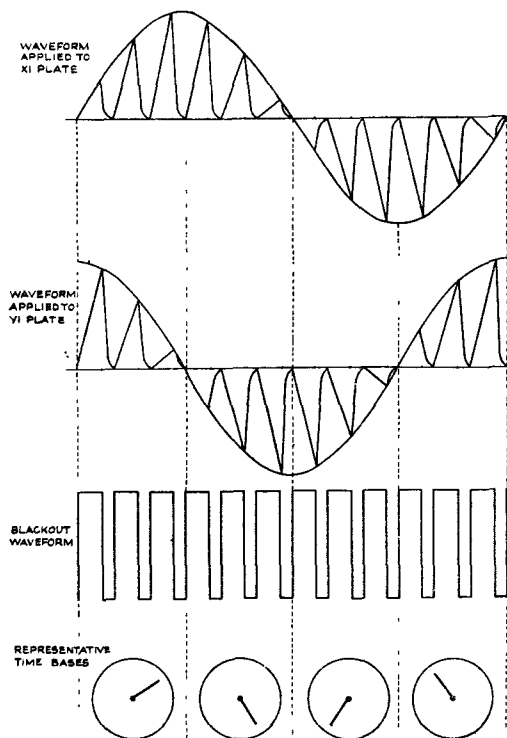


Fig. 186.—Time base waveforms for electrostatic PPI

333. A radial time base sweep of the length r along OC, where angle $AOC = \theta$, may be regarded as the resultant of two time base sweeps, one of length $r \cos \theta$ along OA and

the other of length $r \sin \theta$ along OB. Thus if the sweep along OA (or OB) is produced by a voltage sawtooth waveform of amplitude V , then similar waveforms of amplitudes $V \cos \theta$ and $V \sin \theta$ must be applied to the Y1 and X1 plates respectively to produce a sweep of the same length along OC. In fig. 185, θ is an acute angle, but the above results hold for all values of θ , a negative value of $V \cos \theta$ or $V \sin \theta$ indicating that a negative going waveform must be applied to the appropriate plate. Thus, if θ lies between 90 deg. and 180 deg., then $V \cos \theta$ is negative and $V \sin \theta$ positive, so that a negative going sawtooth waveform is required for the Y1 plate and a positive going waveform for the X1 plate.

334. The waveforms required to give a radial time base, which rotates at uniform speed, are sawtooth waveforms with amplitudes varying sinusoidally with time and in quadrature. Fig. 186 shows one complete cycle of the waveforms applied to the Y1 and X1 plates. The bright-up waveform required for application to the grid is a square waveform of the same period as the sawtooth, and phased as shown in the diagram. A representative time base is also shown in each quarter cycle. For the sake of simplicity only a few time base waveforms are shown in the diagram, whereas a time base is started each time the transmitter fires. The H2S, Mark II PPI display has about 670 time bases during each revolution of the scanner, and therefore the angle between successive radial sweeps is about $\frac{1}{2}$ deg.

335. In the above discussion it has been assumed that the sawtooth voltages rise from zero; if this is not so then the time base sweeps will not start from the centre of the screen. The sweep can be made to start from the centre by applying the necessary shift voltages to the plates.

336. The rotating radial time base can be produced as described above by applying sawtooth waveforms to only one deflector plate of each pair, but in order to prevent the types of distortion discussed in para. 255 onwards it is necessary to apply balanced waveforms to each pair of plates.

Magslip

337. A magslip is employed to produce the sinusoidal variations of amplitude of the time base waveforms used in an electrostatic PPI display. The magslip is essentially a transformer having a rotating primary winding (the

rotor) and two stationary secondary windings (the stators). The stators are at right angles to one another as shown in fig. 187, so that when the rotor is making full coupling with one stator it makes zero coupling with the other. The output from the first stator is then at maximum and the output from the second is zero. Suppose that the rotor lies

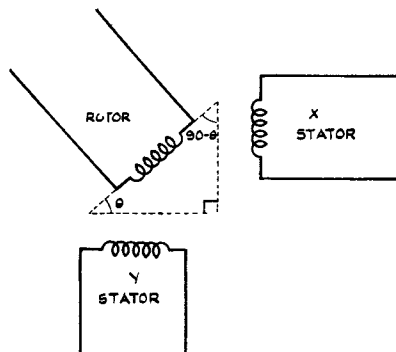


Fig. 187.—Magslip

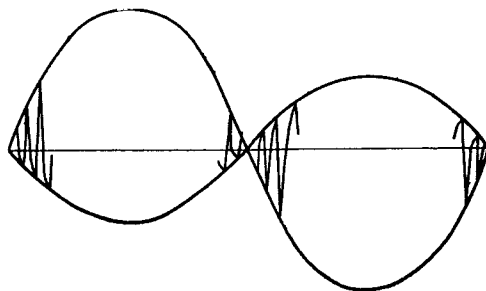


Fig. 188.—Output from stator

so that its axis makes an angle θ with the axis of the stator feeding the Y plates and an angle $(90 - \theta)$ with the X stator. Now if a sawtooth voltage waveform is fed to the rotor, then primary current will flow and produce a magnetic field varying with the input voltage. Since the flux linkage of this field with the Y and X stators is in the ratio $\cos \theta : \sin \theta$, then sawtooth voltages with amplitudes in this same ratio will be induced across the stators. By driving the rotor, in synchronism with the aerial system or scanner, sawtooth outputs with amplitudes varying sinusoidally with time but in quadrature are obtained from the stators. Synchronism with the scanner may be achieved either by direct gearing of the rotor to the scanner, or electro-mechanically by using selsyn motors.

338. The output from one stator of a magstrip is shown in fig. 188. Note that the mean level of each time base is at zero. This means that the time base voltage does not start from the zero level, and so the trace does not start from the centre of the screen. Several ways of ensuring that the trace starts from the centre have been devised, but only the method employing clamping circuits will be discussed in detail. The stator outputs are clamped by a special type of D.C. restoration circuit, each time base sweep being made to start from the same voltage level

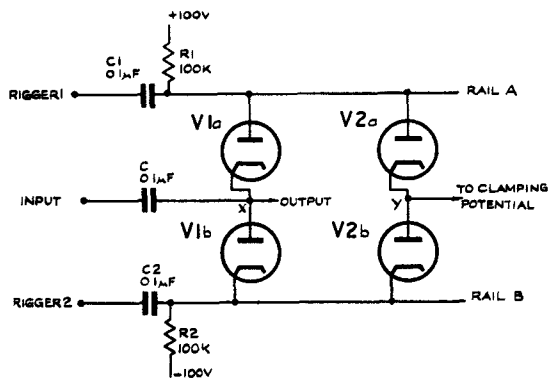


Fig. 189.—Diode clamping circuit

Diode clamps

339. Fig. 189 shows a circuit using two double diodes, e.g. VR54, which will clamp the output of one stator to any desired potential. Anti-phase trigger square waveforms, related to the time base as shown in fig. 190, are fed via the long CR circuits C1R1 and C2R2 to the rails A and B respectively, R1 being returned to +100 volts and R2 to -100 volts. The point V is held at the clamping potential, which is zero if the trace is required to start from the centre of the screen. The diodes V2a and V2b DC restore the trigger waveforms so that during the flyback period, rail A is held slightly positive and rail B slightly negative with respect to the clamping potential, and during the forward stroke rail A is well negative, rail B well positive. Thus all the diodes are conducting during the flyback, but are cut off for the period of the forward stroke.

340. The time base output from the stator is applied through the capacitor C to the point X, from which the clamped output is taken. Assuming that the circuit has reached a steady state by the end of the flyback period, the potential at X is the mean of the two rail potentials, i.e. approximately the clamping

potential. Since the diodes are cut off throughout the forward stroke, the voltage variation at X is the same as that of the input to the capacitor C, but starts from the clamping potential. At the end of the forward stroke the diodes are brought into conduction by the trigger waveform, and the rails A and B are again brought to potentials approximating to the clamping potential. The capacitor C rapidly discharges through the diodes V1a and V1b, so that the point X quickly returns to the clamping potential. The circuit has now returned to the initial conditions; the above operation is repeated for successive cycles of the time base waveform.

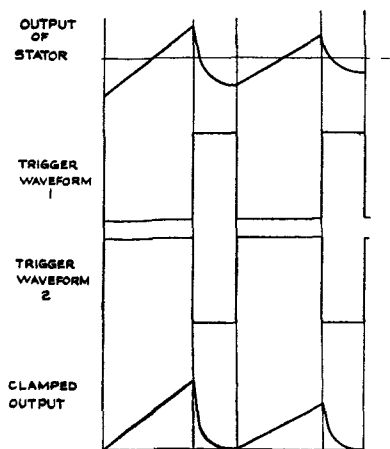


Fig. 190.—Waveforms for diode clamping circuit

341. The amplitude of the trigger waveforms must be greater than that of the time base sweep, to ensure that neither of the diodes V1a and V1b is brought into conduction during the forward stroke. Thus for the waveforms shown in fig. 190, the anode of V1b is being taken positive by the time base sweep, say from zero to +100 volts, and so the trigger waveform should take rail B to a potential higher than +100 volts, in order to hold the cathode voltage positive with respect to the anode.

342. By duplication of the circuit consisting of the capacitor C and double diode V1, both stator outputs may be clamped simultaneously to the same voltage.

Triode clamps

343. The clamping action described above can also be produced by the circuit shown in fig. 191. A double triode such as the 6SN7 is

often employed in this circuit. The waveforms are similar to those shown in fig. 190, ignoring the second trigger waveform.

344. A trigger square waveform with its negative going portion corresponding to the forward stroke of the time base is applied to the grids of V1 and V2 via the coupling circuit C1R1, R1 and the cathode of V2 being returned to the clamping potential. DC restoration occurs due to the diode action of the grid-cathode space of V2, so that both grids are held at the clamping potential during the flyback, and well negative during the forward stroke. Thus if the circuit reaches a steady state by the end of the flyback period, then the anode current flowing in V2 (the grid-cathode voltage of which is zero) must all be drawn through V1. To allow current to flow in V1 its grid can only be slightly negative with respect to its cathode, i.e. the point Y. Consequently, the potential of point Y is very nearly the clamping potential.

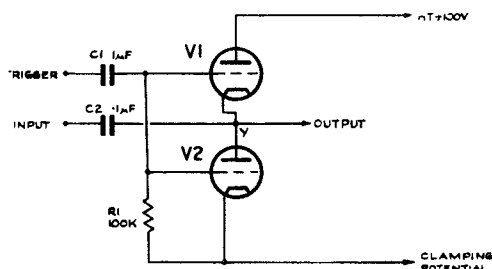


Fig. 191.—Triode clamping circuit

345. The time base output from the stator of the magflip is fed through the capacitor C2 to the point Y, from which the output is taken. The amplitude of the trigger waveform is made great enough to cut off the valves during the forward stroke, and so during this period the voltage variation at Y is the same as the output from the stator, except that it starts from the clamping potential.

346. At the end of the forward stroke the triodes are brought into conduction, and C2 discharges through the valves, so that Y rapidly returns to the clamping potential. This operation is repeated for successive time base cycles.

"Balanced" time base waveform

347. Some of the newer radar equipments dispense with clamping circuits by employing a "balanced" time base waveform as shown in

fig. 192. The time base has an overswing flyback which is so controlled as to make the forward stroke start from the mean level. When this time base is fed to the rotor of the magflip, the forward stroke of the output from either stator will always start from zero, and therefore no clamping is required. The circuits employed are rather complex and so will not be discussed.

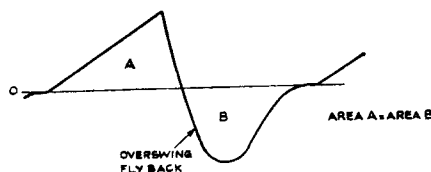


Fig. 192.—Balanced time base waveform

Electromagnetically deflected PPI

348. Electromagnetic CRTs give a better brilliance and focus than electrostatic tubes, but heavy currents (50–100 mA) are required in the coils to give the requisite deflections. This disadvantage has tended to restrict the use of electromagnetic tubes to ground equipments, although increasing use of such tubes is now being made in airborne equipments.

349. Two methods of obtaining a rotating radial time base for an electromagnetic CRT are described below.

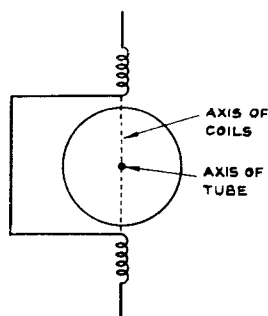


Fig. 193.—Mounting of single pair of deflector coils

(i) The first method employs only one pair of deflector coils; the required time base and shift currents (para. 314) are passed through the deflector coils, which are mounted with their common axis perpendicular to the axis of the tube as shown in fig. 193. The radial time base is rotated by revolving the deflector

coils around the neck of the tube in synchronism with the scanner. Fig. 194 shows the time bases produced for three different positions of the coils.

(ii) In the second method the required currents flow in two pairs of deflector coils which have their axes perpendicular to one another and to the axis of the tube. The problem is very similar to that of the electrostatically deflected PPI where time base voltages are applied to two pairs of deflector plates. Thus, to produce a time base inclined at an angle θ to the axis of one pair of coils, a current $I \sin \theta$ must flow in this pair and a current $I \cos \theta$ in the other pair of coils. The time base is made to rotate by applying current time base waveforms with amplitudes varying sinusoidally and in quadrature to the two pairs of coils.

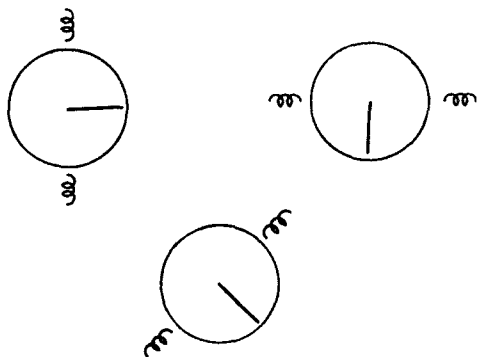


Fig. 194.—Time bases produced for different positions of deflector coils

350. As in the case of the electrostatic PPI it is necessary that the currents flowing in both pairs of deflector coils start from zero so that the time base starts from the centre of the screen. Since it is impracticable to use clamping circuits for current waveforms “balanced” waveforms are employed, i.e. current waveforms of the same form as the voltage waveforms shown in fig. 192. The circuits required to produce these “balanced” current waveforms will not be discussed since they are extremely complex.

CALIBRATION

351. In CRT displays using a linear time base and indicating the reception of a signal by deflection or intensity modulation, an estimate of the range of a target can be made. However, accurate measurement of range can only be made if some scale is provided, graduated in intervals of time. It is convenient to calibrate this scale so that ranges can be read directly.

352. A time base may be calibrated either by deflecting or brightening the spot at the required time instants as in fig. 195. In some displays signals and calibration markers appear on the screen simultaneously, so that the range of a target can be estimated directly from the position of the signal response relative to the calibration markers. In other displays the calibration markers and signal responses can only appear separately on the screen, and it is usual to employ a graduated range scale fixed on the tube face. The time base controls are adjusted when the calibration is switched on so that the calibration markers are lined up against the corresponding graduations of the range scale. It can be seen that calibration necessitates the production of short pulses (or pips) at constant time intervals.



Fig. 195.—Calibrator displays

353. Example

Calibration markers are required at 5 mile intervals.

The velocity of electromagnetic waves is 186,000 miles per second.

The time taken for an electromagnetic wave to travel 5 miles out and 5 miles back is therefore

$$2 \times \frac{5}{186,000} \text{ seconds} \\ = \frac{1}{18.6} \text{ milliseconds}$$

The required pulse repetition frequency is thus 18.6 Kc/s.

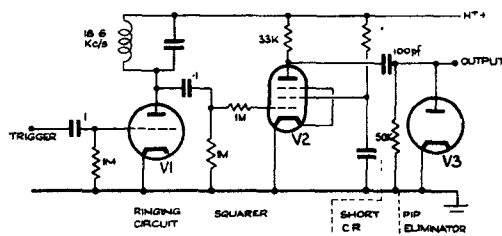


Fig. 196.—Calibrator circuit

Simple calibrator

354. Fig. 196 shows a simple calibrator, and fig. 197 the waveforms obtained at various points. The time base trigger square waveform switches on and off the triode V1 and this shock excites the tuned circuit into oscillation.

The resonant frequency of the ringing circuit is made equal to the required pulse frequency (18.6 Kc/s for the above example). The damped oscillations produced at the anode of V1 are then fed to a squarer pentode V2, biased at zero volts so that symmetrical square waves are produced. Differentiation of the square waves by the short CR circuit produces positive and negative pips, but the diode V3 eliminates the positive pips. The final output is a series of negative pips at the required p.r.f.

355. Note the use of a triode for V1 so that the ringing during the flyback is heavily damped, for V1 is conducting and introduces a low resistance, R_a , in parallel with the tuned circuit. This ensures that there is no residual oscillation at the start of the forward stroke, and so the ring is locked to the start of the time base.

356. The simple circuit described fails to produce a calibration pip at the start of the time base, but this can be overcome by a little modification.

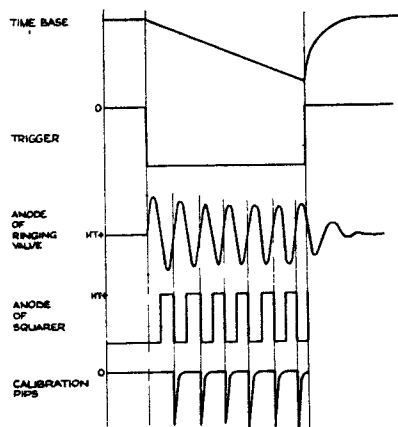


Fig. 197.—Calibrator waveforms

STROBE CIRCUITS

358. Some displays incorporate a device known as a strobe, which enables the operator to select any required portion of the time base. The operator may need the strobe just to "mark" the position of a particular target, e.g. by lining up a bright spot against the front edge of the pulse, or by depressing a small portion of the time base centred about the required signal. Sometimes more detailed examination of this portion of the time base is required, and then a strobe time base is employed to give an enlarged picture of the selected time interval, either on the same or another cathode ray tube.

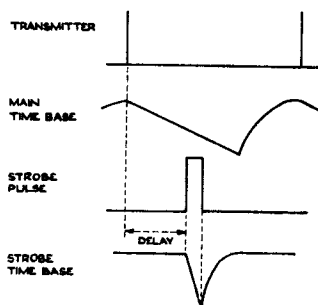


Fig. 198.—Strobe waveforms

359. In order to select a portion of the time base, it is usual to generate a strobe pulse, i.e. a pulse which occurs at some time after the start of the main time base, the delay being controlled by the operator. When an expanded picture is required, this pulse is used to trigger a strobe time base circuit. The relation of the strobe pulse and time base to the main time base waveform are shown in fig. 198.

Strobe circuit

360. A block diagram of a simple strobe circuit and its waveforms are shown in fig. 199. A short positive pulse with leading edge coincident with the firing of the transmitter triggers the flip-flop. The output from the second anode is a square waveform, and this provides the variable delay required for the strobe. Differentiation is produced by a short CR circuit; the strobe pulse is then generated from the differentiated waveform by the pip eliminator and pulse shaper stage. The strobe time base is obtained by using the strobe pulse to suppressor trigger a Miller time base circuit.

Crystal controlled calibrators

357. If precision calibration is required then crystal controlled oscillators are used. The frequency of such an oscillator is generally higher than is required for calibrators; so the crystal oscillators are made to control circuits known as dividers. These give output frequencies which are sub-multiples of the input frequency.

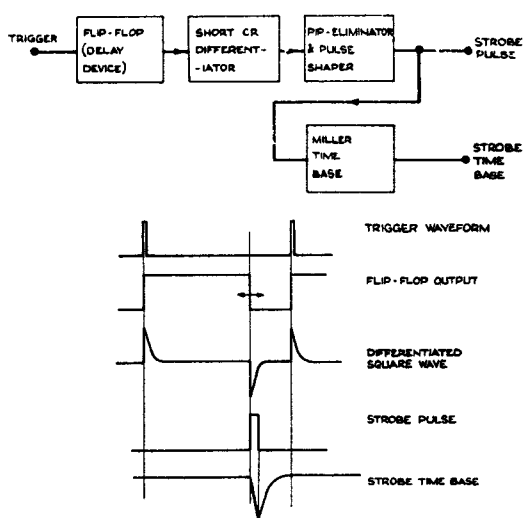


Fig. 199.—Strobe circuit and waveforms

THE SUPER-REGENERATIVE RECEIVER

361. The super-regenerative receiver or detector, often known as the quench receiver, found for many years its main applications in amateur radio circles. It was introduced into certain radar equipments, such as beacons and IFF (Identification Friend or Foe) sets, where portability was required and quality could be sacrificed. Very high amplification (up to 50,000 times) is obtained at frequencies where amplification by other means is almost unattainable; and this is achieved with a very small number of valves. The frequency of operation in radar applications is of the order of 200 Mc/s.

362. The super-regenerative detector is a valve oscillator operating at the same radio frequency as the required signals. The term super-regeneration refers to the employment of an unstable circuit; quenching is the rapid switching of this circuit between an oscillatory and non-oscillatory state. Various methods of quenching and the choice of the quench frequency will be discussed later. It is sufficient for the moment to state that the quench frequency is much less than the radio frequency. The signal is fed from the aerial into the tuned circuit of the oscillator. During the periods in which the detector is oscillatory, oscillations of large amplitude are built up in the normal manner. That is to say, some form of initial disturbance sets the process in operation, and oscillations commence to build up, until the amplitude of the oscillations

reaches some steady value. The time taken to reach this steady state depends upon the strength of the initiating disturbance, being short for a relatively strong disturbance such as the received signal, but longer when the build up is started by the very slight noise disturbance present in all circuits. Thus the power available from the detector depends upon the signal strength at the start of each period of oscillation. The extremely high amplification associated with the super-regenerative detector is obtained because an extremely small signal controls the energy of each burst of oscillation.

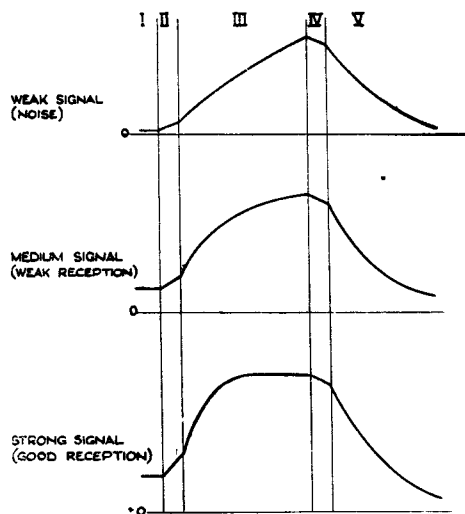


Fig. 200.—Build-up and decay of oscillations in super-regenerative detector

363. The operation of the super-regenerative detector will be examined stage by stage, as shown in fig. 200, which is a qualitative rather than a quantitative diagram.

364. In stage I, the valve is non-conducting and therefore the circuit is in a non-oscillatory state. Nevertheless, there is still a very small oscillation in the tuned circuit picked up from the aerial. Three cases are shown; namely where this oscillation is due to (i) noise, (ii) a weak received signal, and (iii) a strong received signal. It is assumed in (ii) and (iii) that the detector is tuned to the received signal. If this is so then quite a strong oscillation (relative to the noise level) is developed in the tuned circuit; but if the signal is not in tune then the voltage developed is negligible.

365. Stage II is a transition period in which the valve has been brought into conduction,

but not sufficiently to allow self-oscillation. Nevertheless the valve does provide some regeneration, causing some build up of the oscillations picked up from the aerial, though this is only appreciable if the signal is in tune. In fact this stage plays an important part in determining the selectivity of the receiver, and should not take place too rapidly, otherwise the build up of oscillation is negligible.

366. The circuit is self-oscillatory in stage III, and a strong oscillation is built up from the level attained at the end of stage II. With suitable quenching, noise builds up to a fluctuating amplitude, whilst the signal attains a larger amplitude. Two modes of operation are possible according to the strength of the received signal and the length of the period of oscillation. If the amplitude of oscillation does not reach a maximum or saturation level, then the mode is said to be linear; but if saturation occurs then the mode is termed logarithmic, and this gives an automatic volume control action (A.V.C.). The fairly high build up of noise when no signal is present is a noticeable characteristic of the super-regenerative receiver.

367. Stage IV is, like stage II, a transition period in which the valve is conducting, but not sufficient to keep up the oscillation. The amplitude of oscillation therefore decays slightly, but the effect is of little importance.

368. In stage V the valve is not conducting and the oscillations die away fairly rapidly, since tuned circuits of high Q are unattainable at the high frequencies used. Usually the length of this period is sufficient to allow the amplitude to decay to the level being picked up from the aerial, before the next period of valve conduction commences. Then the cycle of operations repeats.

369. As in all receivers, it is now necessary to extract the information contained in the modulated radio frequency carrier signal, that is a second detector is required. Fig. 201 shows the waveforms obtained when the modulation takes the form of short bursts of radio frequency waves; again of necessity in qualitative rather than quantitative form. The output from the super-regenerative stage is a series of short bursts, at intervals depending upon the quench frequency. The second detector, probably a diode, eliminates the radio frequency, but as shown there is still evidence of the quench frequency. This is unavoidable because of the need of a video range of frequencies. The second detector is followed by video amplifier and output stages.

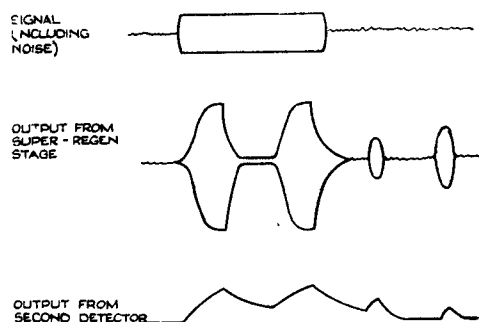


Fig. 201.—Detection of output from super-regenerative stage

370. It is now appropriate to consider quenching in more detail. The general method is to cause the valve in an oscillator circuit to switch on and off; this was assumed in the stage by stage description of the operation of the circuit. This switching may be inherent in the oscillator itself, that is the oscillator is self-quenching. The operation of such a “squegging” oscillator is very similar to that of a blocking oscillator. Radar super-regenerative receivers usually employ a separate quench oscillator; an obvious and common method of switching the valve providing regeneration is to feed on to its grid the output of the quench oscillator. The amplitude of the quench should be of the order of 50 volts. The quench waveform and frequency are determined from the following considerations:—

(i) If the quench frequency is too low, then the output from the second detector has a bad shape, because the quench frequency gets through the detector with ease.

(ii) Stage II must not be too short, as previously explained. This consideration puts an upper limit to the quench frequency, and also means that the waveform should not have a very fast switching edge during stage II.

(iii) If the logarithmic mode is desired, then stage III must be of sufficient length; this also imposes an upper limit on the quench frequency.

371. *Example*

For reception of 5 μ sec. pulses at a radio frequency of 200 Mc/s, a suitable quench frequency is 300 Kc/s. A sine waveform at this frequency is easily obtained; and the bias arrangements should be such that the radio frequency stage is self-oscillatory for about

one-third of the quench cycle. Fig. 202 shows the type of display obtained under such conditions. Quench is getting through the second detector giving a "mush" on top of the pulse, and the "grass" growing on the baseline is due to the build up of noise.



Fig. 202.—Quench mush

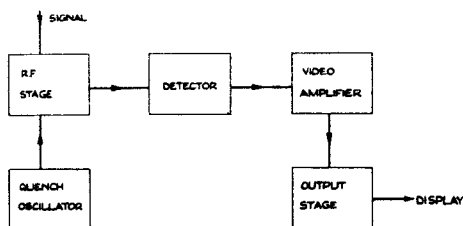


Fig. 203.—Block diagram of super-regenerative receiver

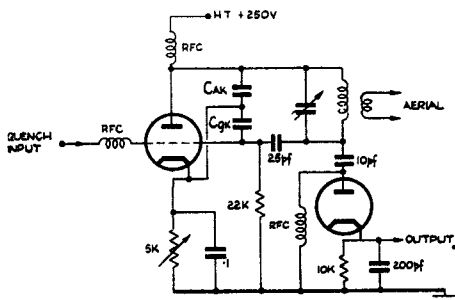


Fig. 204.—RF stage and detector

372. The block diagram of a super-regenerative receiver is shown in fig. 203, whilst fig. 204 shows a possible circuit for the radio frequency and detector stages. The oscillator shown is of the ultra-audion type, i.e. uses the interelectrode capacitance of the valve, and includes a small tuning capacitor. The detector is fed through a small capacitor, and is a diode with a load of appropriate time constant. The other stages are not shown in detail, but may be assumed to be of normal design.

PULSE MODULATION OF TRANSMITTERS

373. The function of a radar transmitter is the production of very short pulses of high frequency radiation at regular intervals of time. Considerable power is required for the

duration of the pulses, but because radiation only occurs for a small fraction of the total time the average power is quite low.

374. Example

It is required to find the average power of a transmitter of peak power 200 KWatts, giving 1 μsec. pulses at a p.r.f. of 400 c/s.

The power during a pulse=200 KW.

The period of a 400 c/s waveform is 2,500 μsec., therefore the average power

$$= \frac{1}{2,500} \times 200 \text{ KW.}$$

$$= 80 \text{ watts.}$$

Now a mean power of 80 watts may be obtained with reasonable heat dissipation, but the peak power is determined by the available cathode emission and the breakdown voltages of the oscillator.

375. The type of oscillator employed for the transmitter varies with the frequency of operation.

20–60 Mc/s conventional valve oscillators.

100–600 Mc/s tuned lines with conventional type valve, modified for V.H.F. working.

3,000 Mc/s and above, magnetrons and klystrons.

376. Pulse modulation is achieved by switching on the oscillator for a time equal to the required pulse length, and by switching the oscillator off until the start of the next pulse. The transmitter therefore radiates short bursts of oscillations at the required p.r.f. If a triode oscillator is used, modulation may take two forms:—

(i) grid modulation, by applying a positive pulse to the grid of the triode.

(ii) anode modulation, by switching the H.T. supply to the oscillator.

There are many disadvantages associated with grid modulation, and so anode modulation is more often used. In fact for magnetron oscillators grid modulation cannot be used because of the absence of a grid. The anode is earthed, and "anode" modulation is achieved by switching the negative voltage supply to the cathode of the magnetron.

377. Grid modulation will not be discussed; anode modulation will be further described for the two forms:—

(i) Series switch modulation.

(ii) Spark gap modulation in conjunction with delay networks. Thyratrons (gas-filled triodes may be used instead of spark gaps).

Series switch modulation

378. Fig. 205 shows the basic circuit of a series switch modulator. The tetrode modulator valve is normally cut off by the negative voltage to which the grid leak resistance R_g is returned; there is then no high tension supply across the oscillator. A positive trigger pulse drives the grid of the modulator positive with respect to the cathode. The modulator valve then conducts hard; most of the high tension voltage is dropped across the oscillator, and only a small fraction of the high tension voltage is taken by the modulator.

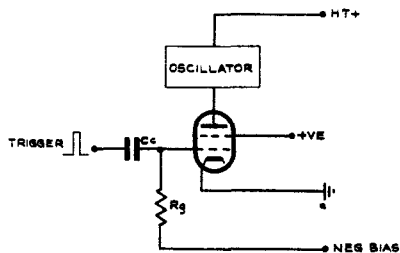


Fig. 205.—Series switch modulation

379. Very high voltages are in use, and the modulator valve must be capable of bearing these voltages without breakdown; a wide electrode spacing is thus indicated. On the other hand during the pulse the modulator should have a relatively low voltage across it, and yet should be able to pass the heavy current of the oscillator, this consideration points to fairly close electrode spacing. Another factor governing the electrode spacing is the maximum interelectrode capacitance which can be tolerated, since such capacitance slows down the edges of the trigger pulse, and consequently the edges of the transmitter output pulse.

380. Tetrodes are used since they have a small input capacitance and yet provide a heavy current. Two or more valves may be used in parallel, but even so the heavy loading of these valves makes very careful design and handling a necessity. The advantage of a hard valve switch is its ability to operate at high pulse repetition frequencies. A disadvantage is the high voltage supply required; also much of the power provided is lost by heat dissipation in the modulator valves.

Pulse transformers

381. It is often necessary to pass a pulse through a transformer, and by careful design

of the transformer this can be done without too much distortion of the pulse. Pulse transformers may be used:—

- (i) for matching purposes and for stepping voltages up or down,
- (ii) for phase reversal.

Use of pulse transformers in conjunction with hard valve modulators

382. In order to distribute the weight equally in the aircraft, it is often necessary to separate the oscillator and modulator valve circuits of airborne equipments, the oscillator for example being placed in the nose close to the scanner and the modulator further down the fuselage. The two units are connected by a cable, which must be terminated in its characteristic impedance to ensure that the pulse shape is not unduly distorted. In the circuit shown in fig. 206 pulse transformers are used to terminate the cable correctly. T1 is used to match the impedance of the oscillator when conducting to the cable. The matched cable will not in general present a suitable impedance for direct connection to the modulator, and so T2 is inserted at the input end of the cable. The characteristic impedance of the cable is normally less than the impedance of both the modulator and the oscillator when conducting; T1 is therefore a step-up and T2 a step-down transformer. Consequently the voltage across the cable is very much less than that applied to the oscillator.

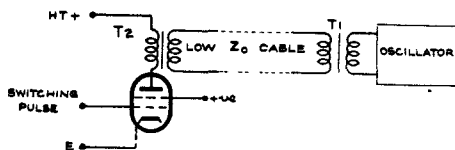


Fig. 206.—Use of pulse transformer

383. Example

Impedance of a magnetron oscillator when conducting = 900 ohms.

Characteristic impedance of cable = 100 ohms.
Amplitude of pulse across magnetron = 12 KV
Therefore step-up ratio of T1

$$= \sqrt{\frac{900}{100}}$$

$$= 3$$

$$\text{Voltage across cable} = \frac{1}{3} \times 12 \text{ KV}$$

$$= 4 \text{ KV.}$$

384. It is usual when a pulse transformer is used for matching purposes to use an auto transformer, i.e. a tapped inductance. This has two advantages:—

- (i) an auto transformer is less bulky,
- (ii) insulation is easier.

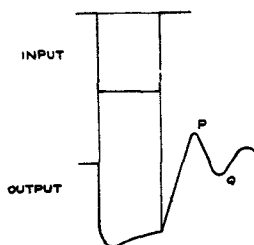


Fig. 207.—Distortion produced by pulse transformer

Overswing diode

385. Even with the most careful design some distortion is produced by pulse transformers. Fig. 207 shows the more important features of the output pulse obtained for a perfect input pulse; the back edge is shown as a simple damped ring, but in practice may be rather more complex. If such a pulse is applied to a magnetron, oscillations occur during the main pulse; there is also the possibility of conduction during the peak Q. This difficulty can be overcome by placing a diode with protective resistor across the magnetron as in fig. 208. The diode conducts during the positive excursion P, and so damps this over-swing to a considerable extent. This reduction of amplitude of the peak P ensures that the amplitude of the following half-cycle is too small to cause the magnetron to conduct.

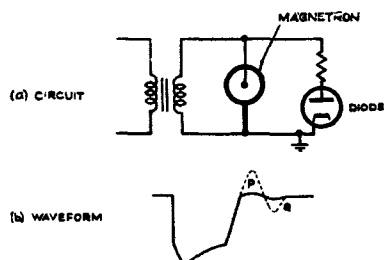


Fig. 208.—Overswing diode

Driver stage

386. This unit provides the positive pulse required to trigger a hard valve modulator. The voltage amplitude of this pulse may be as much as 1,000 volts, for not only does the grid base of a modulator valve approach this

figure, but also the grid must be sent well positive to produce the maximum attainable valve current. Heavy grid current flows, and so the driver must have a low output impedance for this large positive excursion to be achieved. A further advantage of this low output impedance is the rapid charge and discharge of the input capacitance of the modulator, so that there is little loss of steepness on the edges of the trigger pulse applied to the grid of the modulator.

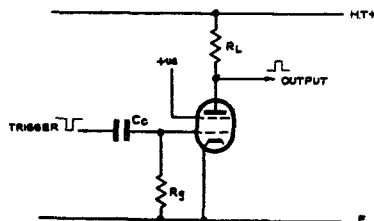


Fig. 209.—Possible driver stage

387. Fig. 209 shows a possible driver using a heavy current tetrode, triggered by a negative pulse obtained by one of the methods described in para. 195 on. Thus for most of the triggering cycle the grid-cathode voltage is approximately zero and the tetrode conducts hard, but for the duration of the pulse the valve is cut off. The output from the anode is therefore a positive pulse suitable for triggering the modulator valve. Since conduction occurs practically throughout the cycle the valve dissipation is excessive.

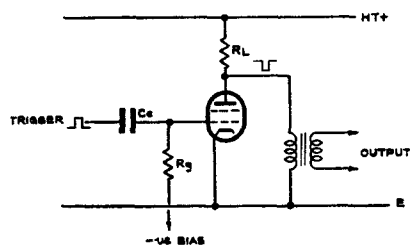


Fig. 210.—Driver stage employing pulse transformer

388. The valve dissipation can be reduced by using the phase reversing property of a pulse transformer, as shown in fig. 210. The driver valve is now normally cut off, but is brought into heavy conduction by a positive trigger pulse. The anode voltage is therefore a negative going pulse, fed to the primary of a pulse transformer. A positive pulse output is taken from the secondary. The transformer may also be used to step up the voltage amplitude of the pulse.

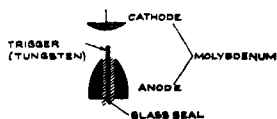


Fig. 211.—Trigratron

Use of spark gaps for modulation

389. The second method of anode modulation employs a spark gap as a switch for the high tension supply to the oscillator. A triggered spark gap, or trigratron, has been developed for this purpose; its electrode structure is shown in fig. 211. The two main electrodes are termed cathode and anode; a third electrode, the trigger, projects slightly from the anode. The electrode assembly is about 1 inch long, and is enclosed in a glass envelope about 4 inches high. A mixture of 97 per cent. argon and 3 per cent. oxygen at a pressure of 3 atmospheres is enclosed in the envelope; this slight amount of oxygen is introduced to give rapid de-ionisation of the gas.

390. During the quiescent period the anode is at earth potential, the trigger at about +300 volts, and the cathode at about -4 KV (this voltage provides the negative high tension supply for the oscillator). These voltages are insufficient to break down the gap. The gas is ionised by sending the trigger electrode to a high positive potential, say 12 KV, and the gap is arced. The opening of the gap is more difficult, and will be described later.

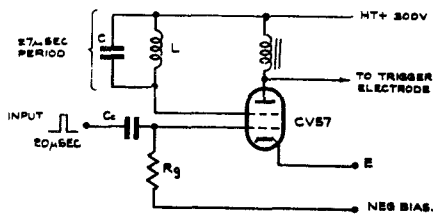


Fig. 212.—Trigger circuit for trigratron

Trigger circuit for trigratron

391. A trigger circuit for applying a quick voltage surge to the trigger electrode is shown in fig. 212, and the waveforms produced in fig. 213. A hard valve of the modulator type, e.g. a CV57 tetrode, is normally biased beyond cut-off by returning its grid to a negative potential, but a 20 μsec. positive pulse input to the grid allows heavy conduction for the period of the pulse. The leading edge of this

pulse shocks the tuned circuit on the screen grid into oscillation. The period of the ring is made 27 μsecs. so that after 20 μsecs. three-quarters of a complete oscillation have been completed and the screen voltage is well above the high tension level. The choke in the anode circuit causes a sudden drop of anode voltage when anode current is switched on by the front edge of the 20 μsec. pulse. After 20 μsecs. a heavy anode current is flowing because of the high screen voltage; and therefore when the valve is cut off by the trailing edge of the input pulse the back e.m.f. of the choke causes the anode to rise to a very high voltage, usually about 12 KV, this being sufficient to trigger the spark gap.

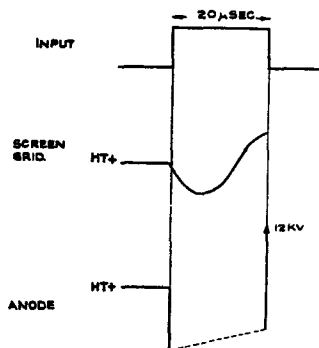


Fig. 213.—Waveforms of trigger circuit for trigratron

Use of delay network to open spark gap

392. A voltage surge applied to the input terminals of a delay network is transmitted along the network, arriving at the far end after, say, a time T . If the network has an open-circuit at this end, then the voltage surge is reflected in phase without loss of amplitude, and then travels back along the network. Thus after a total time of $2T$ a second voltage surge appears at the input terminals, the two surges being of equal magnitude and in the same sense.

393. Associated with a delay network is a certain resistance known as the characteristic or surge impedance; this resistance will be denoted by Z_0 . Its two main properties are:—

(i) If a sudden change of the voltage across the input terminals of the network occurs, then the current flowing into the network also changes suddenly, and the surge impedance, Z_0 , is the ratio

$$\frac{\text{change of voltage}}{\text{change of current}}$$

(ii) If a delay network is terminated at either end by its characteristic impedance, Z_0 , then a wave travelling towards that end is completely absorbed by the termination and no reflection occurs.

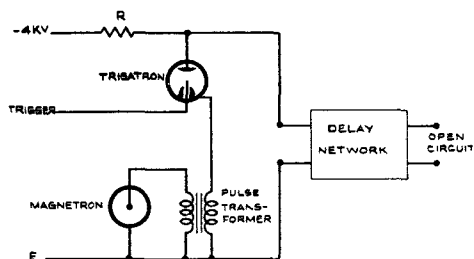


Fig. 214.—Trigatron modulator circuit

394. A modulation circuit using these properties of a delay network is shown in fig. 214. Obviously when the trigatron is open the delay network is charged through the resistance R to a voltage of -4 KV. During this fairly slow charging process the delay network may be regarded as approximately a capacitance C , equal in value to the total capacitance of all the sections in parallel. The circuit for this stage is shown in fig. 215; the charging process is exponential with time constant CR .

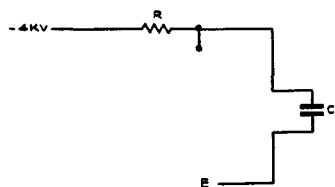


Fig. 215.—Charging circuit

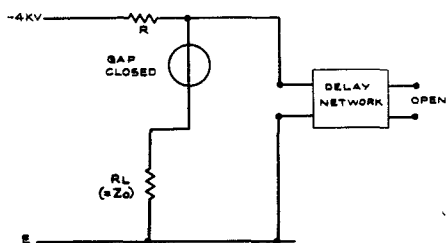


Fig. 216.—Equivalent circuit with trigatron closed

395. Suppose that the delay network has been fully charged when the trigatron is triggered. The gap is arced, and the trigatron may be regarded as a short-circuit. A negative voltage surge appears across the primary of the pulse transformer, and induces a similar surge in the secondary causing the magnetron to conduct.

It is found advisable for optimum working of the circuit to match the magnetron into the delay network, i.e. the turns ratio of the pulse transformer is chosen so that the impedance of the magnetron during conduction, as seen from the primary side, is equal to the characteristic impedance, i.e. $R_L = Z_0$. The equivalent circuit now holding is shown in fig. 216.

396. Typical values of the resistances are:—

$$R_L \text{ and } Z_0 = 100 \text{ ohms.}$$

$$R = 10 \text{ K.}$$

R is considerably larger than the other two values and may be regarded as an infinite resistance or open circuit. The voltage surge across the primary of the pulse transformer is therefore due to the discharge of the delay network into the resistance R_L . Let the value of this surge be $-V$ volts, i.e. the voltage across the oscillator jumps from 0 to $-V$ volts, and the voltage across the network from $-4,000$ to $-V$ volts. The corresponding current jumps are in magnitude,

$$\frac{V}{R_L} \text{ and } \frac{4,000 - V}{Z_0}$$

Now each of these expressions represents the discharge current of the network, and therefore, since $R_L = Z_0$,

$$V = 4,000 - V$$

$$\text{i.e. } V = 2,000$$

Thus the voltage surge across R_L is half of the negative high tension voltage.

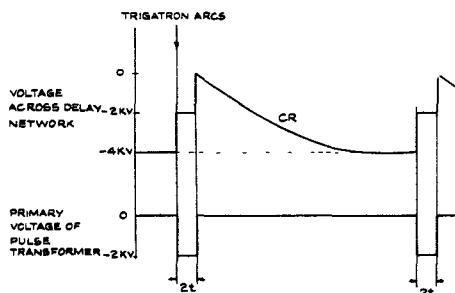


Fig. 217.—Waveforms of trigatron modulator

397. The voltage surge which travels down the delay network is also half of the high tension voltage. Now after time $2T$ this voltage surge arrives back at the input terminals, and has the same amplitude and sense. No further reflection occurs since R_L terminates the delay network in its characteristic impedance. Thus after a time $2T$ the voltage across both R_L and the trigatron have fallen to zero. The trigatron becomes de-ionised, and the gap opens. The

delay network is now charged through R, and the above process repeats when the trigatron is next arced. The waveforms are shown in fig. 217.

398. In practice the turns ratio of the pulse transformer gives a step-up of the 2 KV pulse, and the pulse applied to the magnetron has sufficient amplitude to cause the magnetron to conduct.

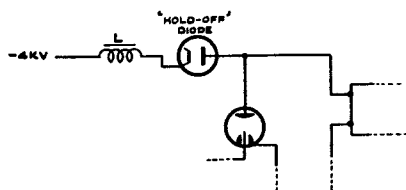


Fig. 218.—Resonant choke charging

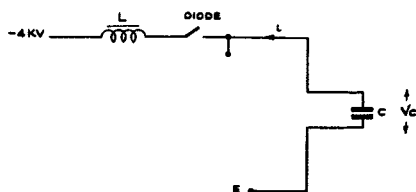


Fig. 219.—Resonant choke charging circuit

Resonant choke charging

399. A serious disadvantage of resistance charging of the delay network is that the amplitude of the pulse produced across the primary of the pulse transformer is only half the applied HT voltage. The pulse amplitude can be made equal to the HT voltage by employing "resonant choke charging". The resistor R is replaced by a choke L in series with a "hold off" diode (fig. 218). The charging circuit may now be represented by fig. 219, the diode, which is shown as a switch, being closed. The charging of C through L is oscillatory and analysis shows that C will charge to twice the applied voltage, i.e. to -8 KV (fig. 220). If the half-period of this oscillation finishes at the instant the trigatron is arced, then the delay network will discharge from -8 KV. The discharge of the network is the same as for the resistance charged line, but for a pulse of amplitude 4 KV is produced across the primary of the pulse transformer.

400. The diode is only required when the line is completely charged some time before the trigatron gap is closed, since in the absence

of the diode, oscillatory discharge of the line through L would occur. This discharge current is in the direction to flow through the diode, and the line is held charged until the trigatron is closed.

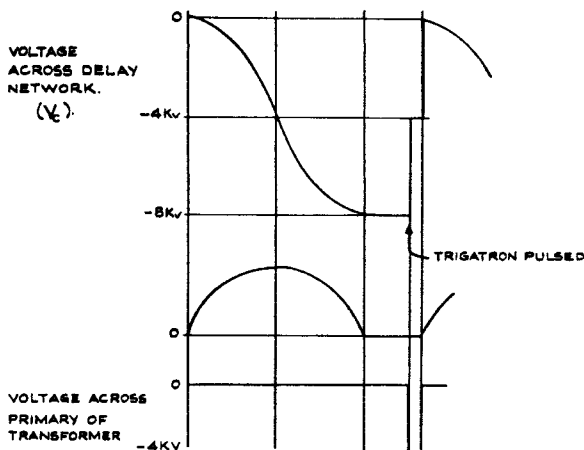


Fig. 220.—Resonant choke charging waveforms

Constant current charging

401. If the time between successive trigger pulses to the trigatron is made less than the half-period of the LC oscillation, then it can be shown that after a few cycles, the delay network will charge from zero to 8 KV in each charging period. In this case, the delay line just completes its charge at the instant when the trigatron gap is closed, the charging current being approximately constant and always flowing in the same direction (fig. 221). For this reason a diode is not included in the circuit when "constant current charging" of the delay network is employed.

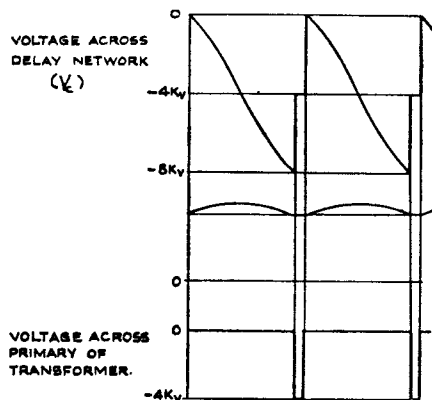


Fig. 221.—Constant current charging

CHAPTER 2

RADAR AERIALS

LIST OF CONTENTS

	<i>Para.</i>		<i>Para.</i>
Introduction	1	Reflection from earth	125
Transmission lines and waveguides	5	Additional phenomena of propagation	145
Simple aerials	16	Limitations to performance of ground equipments	154
Wide band aerials	24	Direction finding	
Reflectors and directors	26	Ground equipments	164
Directional aerials	36	Airborne equipments	175
Yagi aerials	43	Elevation finding	178
Folded aerials	46	Classical method	183
Resonant slots	47	Modern methods	194
Beaming	51	Power gain of aerial systems—	
Broadside arrays	67	mathematical aspects of polar diagrams	200
Paraboloidal reflectors	70	Power carried by electro-magnetic waves	202
Special aerials	98	Factors governing maximum range of radar equipments—	
Cosec θ aerials	99	Noise	211
Polyrod aerials	100	Pulse and band width	222
Waveguide arrays	101	Maximum range of detection	229
Cheese aerials	103		
Effect of ground on radiation patterns—			
Theoretical aspects	107		

LIST OF ILLUSTRATIONS

	<i>Fig.</i>		<i>Fig.</i>
Half-wave aerials	1	Reflection from paraboloids	24
Stacked half-wave aerials	2	Parabolic mirrors	25
Quarter-wave aerial	3	Optimum focal lengths of mirrors	26
Quarter-wave aircraft aerial	4	Feeding circular-aperture mirrors	27
Quarter-wave collar	5	Feeding rectangular apertures	28
Wide-band aerial	6	Beam swinging with paraboloids	29
Wide-band aerial with shorted stub	7	Theoretical polar diagram	30
Evaluation of polar diagrams	8	Cosec θ aerial	31
Axial P.D. of half-wave aerial	9	Waveguide arrays	32
Transverse P.D. of half-wave aerial	10	Cheese aerial	33
Polar diagram of half-wave aerial—perspective	11	Reflection from conducting surfaces	34
Electro-magnetic coupling between aerials ...	12	Lobing due to reflection from conducting surfaces	35
Aerial with tuned reflector or director ...	13	Brewster angle	36
Axial P.D. of half-wave aerial with reflector ...	14	Ground reflection with beamed radiation ...	37
Transverse P.D. of half-wave aerial with reflector ...	15	Lobing due to ground reflection	38
Aperiodic reflector	16	Crossed dipoles	39
Yagi aerial	17	Split-beam method	40
Folded aerial	18	Elevation finding—short ranges	41
Resonant slots	19	Elevation finding—long ranges	42
Uniformly-radiating surface	20	Classical method of elevation finding	43
Axial P.D. for fig. 20	21	Height estimation from performance diagram	44
Broadside array	22	Power/frequency graph	45
Construction of a parabola	23		

CHAPTER 2

RADAR AERIALS

INTRODUCTION

1. In this chapter the chief types of radar aerial systems are surveyed, but the underlying theory of aerials and of transmission lines is not dealt with fully. A fuller introduction to aerial theory is given in Chapter XV of A.P.1093—R.A.F. Signal Manual, Part II (Radio Communication). Transmission lines and waveguides are described more fully in Chapter 3 of A.P.1093F.

2. A radar equipment requires aerials from which to radiate the pulses of high frequency power generated by the transmitter, and also to collect power from the scattered radiations returning from the target. Aerial systems are also responsible for providing measurements of angles of bearing and elevation.

3. The relatively simple appearance of the aerials of radar equipments, as compared with the other constituent units, tends to give a misleading impression of their actual importance. It is largely true, however, that a radar equipment is "designed around" the radiation pattern of its aerial system.

4. In particular, the form of radiation pattern determines the forms of display that can be used as well as the degree of accuracy that can be achieved in measurements of bearings and elevations.

Transmission lines and wave-guides

5. The power from the transmitter has always to be conveyed in some way to the transmitting aerial, and similarly the return signal must be conveyed from the receiving aerial to the receiver. To do this it is necessary to employ either transmission lines or waveguides. Transmission lines are used for the longer wavelengths, but for ultra short waves the attenuation of transmission lines is excessive, and it is necessary to use guides.

6. There are two types of transmission line in common use:—

- (1) twin or balanced line, consisting of two parallel wires separated by a distance which must be small in comparison with the wavelength,
- (2) concentric or unbalanced line; consisting of two coaxial cylinders.

Both types of line act as guides to electromagnetic waves which travel along them with a speed approximating to that of light.

7. Waveguides are simply hollow metal tubes which carry electromagnetic waves in much the same way as speaking tubes carry sound waves. They are usually of either rectangular or circular cross-section. Whereas transmission lines will carry waves of any wavelength, there is an upper limit to the wavelength of the waves that can be carried by a guide. This limit depends on the type of guide used and on the way in which the waves are launched into it, but it is generally true to say that the cross-section of the guide must have dimensions comparable with the length of the waves that it is to carry. Thus for waves longer than about 10 cms., guides are too large and unwieldy, and lines are invariably used.

8. A full treatment of the theory of transmission lines and waveguides is given in Chapter 3 of this publication. As, however, it is necessary to refer to certain properties of feeder systems in this Chapter, a brief outline is given here.

9. The most important parameter of a transmission line is its characteristic impedance, which can be defined as follows. Suppose that any type of transmission line has input terminals A B; and consider the line to extend to infinity. If a fluctuating voltage is applied to the terminals A B, electromagnetic waves will travel along the space surrounding the conductors, and these waves will be accompanied by current and voltage fluctuations in the conductors themselves. It can be shown that the ratio of voltage to current is always constant no matter what the frequency of the applied voltage oscillations may be. In other words, this infinitely long line acts as an impedance.

10. If the line has no loss, that is, if the conductors have no resistance in themselves, there is no dielectric loss due to faulty insulation, and no loss due to radiation. Then the characteristic impedance is a pure resistance, and its value depends only on the dimensions of the line. If, on the other hand, the line has loss, the characteristic impedance is partly

reactive; but since all lines used in radar are virtually loss-less, the reactive part is always so small that it is negligible.

11. When a finite length of line is terminated by a load of some kind, for example, an aerial array, the electromagnetic waves and their associated current and voltage waves will in general be partially reflected from the termination, and only part of the available power will be absorbed in the load. Only one termination will absorb all the incident power: namely, a terminating load which is a pure resistance equal in value to the characteristic impedance of the line. Such a load is said to "match" the line.

12. If the load is not matched into the line, the presence of the reflected waves alters the ratio of voltage to current at the input terminals, so that the input impedance is no longer equal to the characteristic impedance. When a line is used to carry power from a generator to a load, unless the load is equal to the characteristic impedance, it is usually necessary to employ some device to match it artificially so that none of the energy is reflected back. There are various methods of matching loads into lines, details of which can be found in text books and reports on the subject. In radar it is necessary to match the transmitting aerial into its transmission line and also to match the receiver into the line from the receiving aerial.

13. For some purposes it is usual to employ lengths of line which are deliberately mismatched at their termination. Two particular cases of this may be mentioned here since they are used so widely. If a length of line is terminated by a short circuit its input impedance is not necessarily zero, as might at first sight be

expected, but is dependent on the length of the line. If the line is a quarter of a wavelength long its input impedance is, in fact, infinite. If the length is somewhat less than a quarter wavelength the line behaves as a pure inductance and if the length is somewhat greater than a quarter wavelength it behaves as a pure capacitance. A line terminated by an open circuit, on the other hand, acts as a capacitance if it is less than a quarter wavelength long, as a short circuit if it is exactly a quarter wavelength long, and as an inductance if its length is somewhat greater than a quarter wavelength.

14. Many radar equipments use the same aerial array for transmission and reception. This is usually accomplished by spark gap switches, which consist of a suitable network of quarter-wave lines terminated at various points by spark gaps. When the transmitter sends out a pulse a high voltage is developed and the gaps break down, so that the lines are short-circuited at certain points. The network is so constructed that this automatically switches the transmitter through to the aerial and disconnects the receiver. When the transmitter is not radiating the spark gaps are open-circuited, and the aerial remains switched through to the receiver.

15. Most of the facts about lines stated above apply also to guides. Although a guide consists of only a single conductor, it is still possible to define a quantity known as the characteristic impedance, which is a measure of the ratio of the amplitudes of the electric and magnetic fields in the waves travelling down the guide. If the guide is not correctly matched at its termination the waves will be partially reflected. Consequently matching precautions are necessary just as they are with transmission lines.

SIMPLE AERIALS

16. The simplest aerial is the *half-wave aerial* (frequently called a dipole). This is merely a metal rod cut to a resonant length of $\frac{\lambda}{2}$, where λ is the operational wavelength of the radar device. Since, in radar, the wavelengths used are always less than 14 metres, the half-wave aerial is usually of a convenient size. The half-wave aerial may be fed from a transmission line as shown in fig. 1 (a). Here the aerial is bisected; the halves are separated and are attached to terminals of a transmission line which supplies power in the form of current

associated with a relatively small potential difference across the gap. A standing wave is excited on the aerial with a current maximum at the gap and virtually zero current at the ends. The input impedance of the aerial at the gap has a small value, with a resistive component of about 73 ohms and a reactive component of about 42 ohms in series.

17. Alternatively, the aerial is not bisected but fed at one of its ends by attaching it to a point of high oscillating potential in the standing wave pattern on a transmission line. Fig. 1 (b)

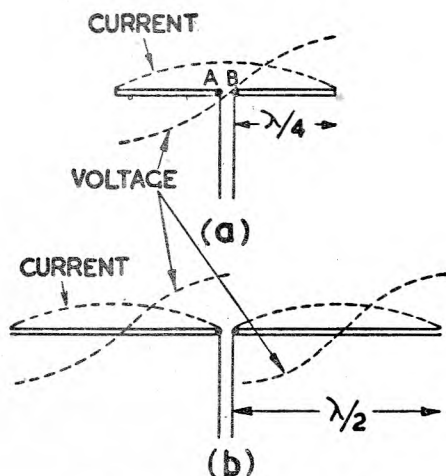


Fig. 1—Half-wave aerials

shows a pair of end-fed half-wave aerials fed from the end of a twin transmission line so as to oscillate in phase. Since the feed-current is small the impedance across the gap is large. Fig. 2 illustrates the excitation of a "stack" of half-wave aerials end-fed from the antinodal potential points of a transmission line.

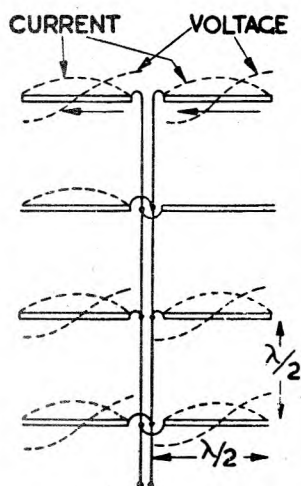


Fig. 2—Stacked half-wave aerials

18. The resistive term in the input impedance arises from the power lost to the system in the radiation emitted by the aerial into the surrounding space. This resistance is, therefore, called the *radiation resistance*.

19. The radiation resistance of a half-wave aerial, centre-fed, is about 73 ohms, and this value does not change markedly with the thickness of the aerial. The radiation resistance of a

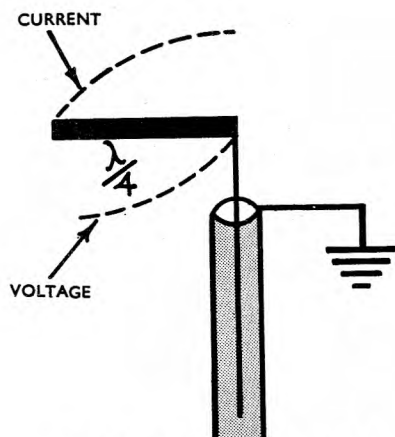


Fig. 3—Quarter-wave aerial

pair of end-fed half-wave aerials is large and increases as the aerials are made thinner. It may have any value up to 10,000 ohms according to the ratio of the aerial diameter to its length. This latter method of feeding is often termed *voltage feeding*, while the previous method is known as *current feeding*. Another method of voltage feeding is that shown in fig. 2. A stack of half wave aerials, spaced at half-wave intervals, are fed from a twin line as shown. A system of aerials such as this is known as an aerial array.

20. In certain aircraft equipments quarter-wave aerials are used. Fig. 3 shows a quarter-wave aerial and indicates the usual method of feeding. An aerial of this type may be regarded as half of a centre-fed half-wave aerial, and the current and voltage amplitudes are shown by the dotted lines. An important difference between the centre-fed aerial and the quarter-wave aerial, however, is that the former has two terminals (A and B in fig. 1 (a)) whose voltages fluctuate in antiphase, whereas the latter has only one terminal. In other words, the half-wave aerial is balanced, while the quarter-wave aerial is unbalanced.

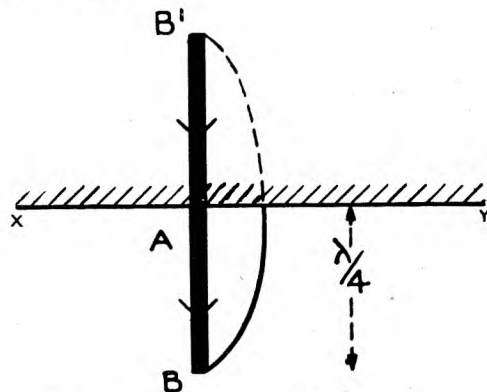


Fig. 4—Quarter-wave aircraft aerial

21. For this reason it is necessary to feed the quarter-wave aerial with an unbalanced line, the outer conductor of which is earthed, while the inner is connected to the aerial. To feed a quarter-wave aerial with a balanced line, it would be necessary either to earth one of the conductors at its termination or to leave it disconnected. In either case the balance of the line would be upset.

22. The quarter-wave aerial is usually mounted in an aircraft as indicated in fig. 4; AB represents the aerial, and XY a conducting sheet which will usually comprise part of the fuselage of the aircraft. Provided that the surface XY is almost plane in the neighbourhood of the aerial, an image AB' of the aerial will be formed by reflection of the waves, and for radiation purposes the arrangement will behave as a half-wave aerial.

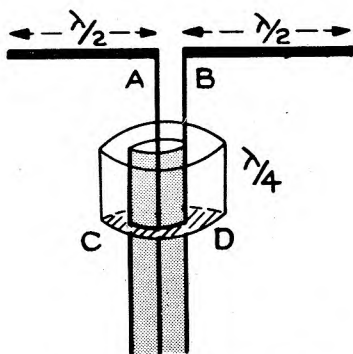


Fig. 5—Quarter-wave collar

23. The comparison of half-wave and quarter-wave aerials is an example of the general rule that a balanced line must be used to feed a balanced load, while an unbalanced load requires an unbalanced line. A second example is shown in fig. 6, which represents a typical aerial used for 10 cm. wavelengths. At this frequency it is possible to use either a waveguide or a line, but if a line is employed it must be coaxial, since a twin line would be too lossy.

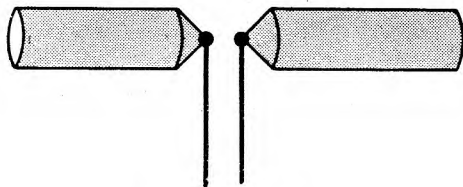


Fig. 6—Wide-band aerial

The problem of feeding a balanced half-wave aerial from an unbalanced line then arises. The inner conductor is connected to the point A on the aerial and the outer conductor to the point B. If the outer conductor is earthed, however, the point B will not fluctuate in potential as it should, while if the outer is left unearthed some of the power will travel down the outside of the line and will be lost. The usual arrangement for overcoming the difficulty is that shown in fig. 7. A quarter-wave collar consisting of a quarter wavelength of tubing of sufficient diameter to slip over the outer of the transmission line, is fixed on to the line making metallic contact at CD. This arrangement effects a balanced-to-unbalanced transformation.

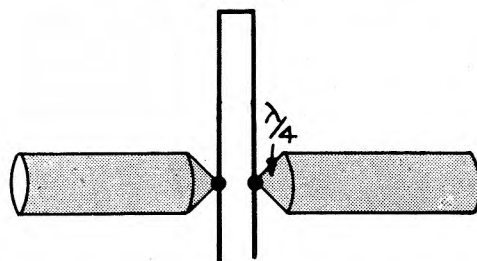


Fig. 7—Wide-band aerial with shorted stub

Wide band aerials

24. It is necessary in certain radar equipments to change the frequency, consequently aerial systems must then be constructed to work over a wide frequency band. This presents difficulties for the following reason. Suppose that an equipment is to work over a certain range of frequencies, and suppose that its radiating element is a half-wave aerial whose length is correct (slightly less than a half wavelength) at the mid-band frequency. If the frequency is increased, so that the wavelength is decreased, the aerial will be too long. Now an aerial slightly greater in length than a half wavelength no longer acts as a pure resistance of 73 ohms, but as an inductive impedance, that is, it behaves as a resistance, somewhat different from 73 ohms, in parallel with an inductance. Thus it will no longer properly match the line, and will not radiate efficiently. If the frequency is decreased, on the other hand, the aerial becomes capacitive and the same trouble arises.

25. There are two methods of overcoming this effect. First, the increase in reactive component of input impedance corresponding to a

given change in frequency is less for aerials of large diameter than for thin aerials; it is better, therefore, to use a very thick rod or tube rather than a thin wire for wide band work. Fig. 6 shows a wide band aerial of this type. Occasionally it may be found that this is not sufficient, however, and that a second method of further increasing the bandwidth of the aerial is also necessary. This is usually accomplished by a device similar to that shown in fig. 7. A length of shorted transmission line, or *shorted stub* as it is often called, is connected across the input terminals of the aerial. Suppose

REFLECTORS AND DIRECTORS

Polar diagram of half-wave aerials

26. The electromagnetic waves radiated from an aerial consist of alternating electric and magnetic fields. Consider a half-wave transmitting aerial whose axis lies in the direction YY^1 in fig. 8, the aerial itself being situated at the point O. A sphere is drawn with the point O as centre, so that all points on its surface are equidistant from the aerial, and its radius is so large that the size of the aerial is negligible in comparison with the dimensions of the drawing.

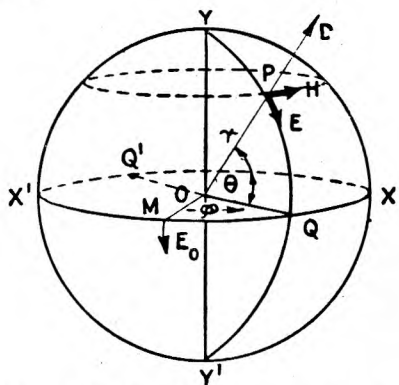


Fig. 8—Evaluation of polar diagrams

27. At any point such as P on the surface of the sphere there will be an electromagnetic field due to radiation from the aerial. If the line YY' is taken as the axis of the sphere, the electric part of this field, E, will be directed along the line of longitude through P, while the magnetic field H will be directed along the line of latitude. As the electromagnetic waves from the aerial sweep past P in the direction of D, the intensity of these fields will fluctuate, and the amplitude of their fluctuations must clearly depend on the distance of P from the aerial. It can, in fact, be shown that the amplitudes of both E and H are inversely proportional to the distance OP.

that the length of this line is exactly a quarter wavelength at mid-band frequency. When the frequency increases, however, the aerial will be too long as before, and will, therefore, become inductive; the length of transmission line, however, will also be greater than a quarter wavelength. Since a length of shorted line somewhat longer than a quarter wavelength acts not as an open circuit, but as a capacitance, it tends to cancel the inductive reactance of the aerial. The opposite happens when the frequency is below mid-band frequency; the aerial becomes capacitive and the line inductive.

The amplitudes depend not only on this distance, however, but also on the direction of P from the aerial. Thus, as P moves along the line of longitude YQY¹ the amplitude of the waves reaching it will be greatest when it is at the point Q, and will fall off towards the poles, until when it is at Y or at Y¹ it will receive no radiation whatever. In other words a half-wave aerial radiates better in certain directions than in others. The directional properties of an aerial can be indicated graphically by a polar diagram such as that shown in fig. 9.

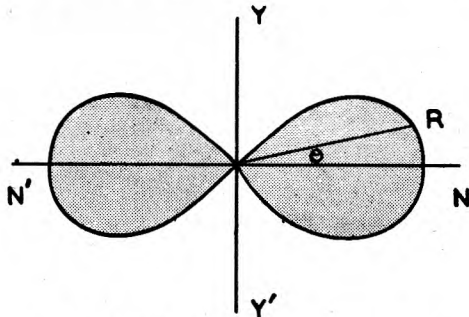


Fig. 9—Axial P.D. of half-wave aerial

28. The line ON in fig. 9 is proportional in length to the amplitude of the waves at the point Q on the equator of the sphere. The amplitude of the waves at P is given by the length of the line OR, where the angle θ between OR and ON is equal to the angle between OP and OQ in fig. 8. Thus fig. 9 is a graphical representation of the relationship between amplitude and direction at constant range.

29. Suppose now that the point P moves down to Q so that the corresponding amplitude is ON, and that it then moves away from Q round the equator of the sphere in the direction QXX¹Q¹, then the line OP will always be perpendicular to the axis of the aerial. A half-wave

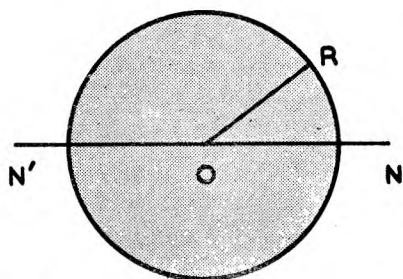


Fig. 10—Transverse P.D. of half-wave aerial

aerial, therefore, radiates equally well in any direction perpendicular to its own length; thus the polar diagram showing the variation of amplitude with direction is now similar to that in fig. 10. The amplitude OR remains constant for all directions, so that the figure is a circle with centre O. Both fig. 9 and fig. 10 are, of course, cross-sections of the complete polar diagram which is a three-dimensional figure, and is shown in fig. 11.

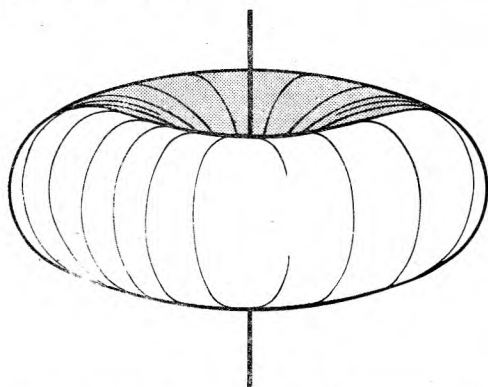


Fig. 11—Polar diagram of half-wave aerial—perspective

30. There is an alternative way of regarding the polar diagram of a transmitting aerial. Suppose that, instead of moving over the surface of a sphere with the aerial as its centre, the point P in fig. 8 moves in space in such a way that the amplitude of the waves reaching it from the aerial remains constant. Since the aerial radiates more energy in certain directions than in others, it is clearly necessary for P to approach O more closely when it is in certain positions than when it is in others, as the polar diagram in fig. 11 then gives a measure of the distance OP. In other words, the polar diagram has two meanings:—

- (1) It measures variation of amplitude with direction at constant range.
- (2) It measures the variation of range with direction for constant amplitude.

31. A half-wave aerial also has directional properties for receiving, since it is more sensitive to radiation falling on it from certain directions than from others. It is possible to indicate these directional properties by a polar diagram just as it was in the case of the transmitting aerial, but the polar diagram now has a different meaning.

32. This meaning can best be understood by considering an imaginary source of electromagnetic waves at a considerable distance from the receiving aerial. Suppose that the aerial at O in fig. 8 is a receiving aerial, and imagine a transmitter situated at P. The radiation from this transmitter will set up a signal in the aerial. As the transmitter is moved over the surface of the sphere the amplitude of the signal in the aerial will vary. The polar diagram is a graph showing the variation of amplitude of signal with direction. It is also possible to move the transmitter in such a way that the amplitude of the signal received remains constant. The same polar diagram will then indicate the variation of distance between the transmitter and receiving aerial with direction for the same signal strength. Once again, then, the polar diagram of a receiving aerial has two alternative meanings:—

- (1) It shows the variation of amplitude of signal received with direction, when the amplitude of the incident waves remain constant.
- (2) It measures the variation of range of the source with direction when the amplitude of the received signal remains constant as a source of electromagnetic waves is moved round the aerial.

33. From the *reciprocity theorem*, for the explanation of which see A.P.1093E, it may be shown that the polar diagram of any aerial has the same shape whether the aerial is used for transmission or for reception. It follows from this theorem that fig. 9 and fig. 10 represent the polar diagram of a half-wave aerial for both transmission and reception.

34. Consider now two half-wave aerials a considerable distance apart, and suppose for simplicity that they are both parallel. If the

aerial at A in fig. 12 is a transmitting aerial, and the aerial at B a receiving aerial, a signal will be set up in B and will be fed into the receiver. The electromagnetic fields due to the waves from A will be propagated from A to B in the way shown, the electric field being parallel to the axes of either aerial, and the magnetic field perpendicular to this direction.

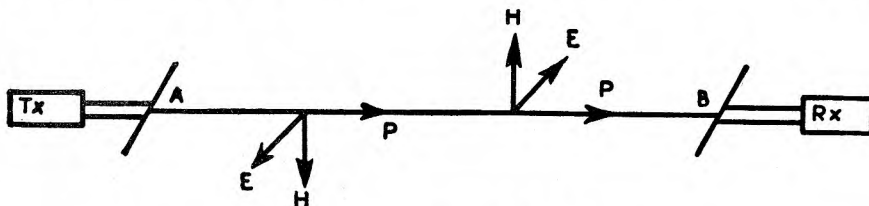


Fig. 12—Electromagnetic coupling between aerials

35. The electromotive force set up in the receiving aerial can be regarded in two ways. It is possible to consider it as being induced by the fluctuating electric field applied along the length of the aerial, and as being induced by the fluctuating magnetic field at right angles to the aerial. It is usually more convenient to think in terms of electric field, however, and to consider the signal in the aerial as being set up

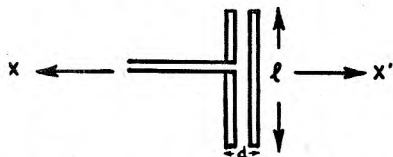


Fig. 13—Aerial with tuned reflector or director

by a fluctuating voltage between its ends, so that the magnetic part of the field is generally neglected for most practical purposes. When the aerial is horizontal with respect to the earth, so that the electric field is directed horizontally, the waves are said to be horizontally polarised, while when the aerial is vertical it is said to give vertically-polarised waves.

Directional aerials

36. It is often necessary to increase the directional properties of a half-wave aerial, and one method of doing this is to use reflectors and directors. Fig. 13 shows a simple reflector or director. A half-wave aerial is driven from a transmitter, and a second aerial is placed a short distance away from it. The distance between the two is usually somewhat less than a quarter wavelength, and is never greater than a half wavelength. The second aerial is not exactly a half wavelength long, but has a length which

may be somewhat greater or less than a half wavelength; it is not connected to a transmitter but carries a high frequency current because of its proximity to the driven aerial. Such an aerial is termed as a *parasite*.

37. It is clear that the amplitude and phase of the current induced in it will depend on the

distance d and the length l , and it can be shown that the length is the more critical of those two factors. The driven aerial and the parasite will each radiate electromagnetic waves, and at any distant point in space the amplitude of the waves received from the system will be the sum of the amplitudes of the separate wave trains for the two aerials. In adding these two amplitudes, however, it is necessary to take into account that the two trains of waves will in general have different phases, and that although they will add to give a large amplitude in certain directions, they will tend to cancel in other directions.

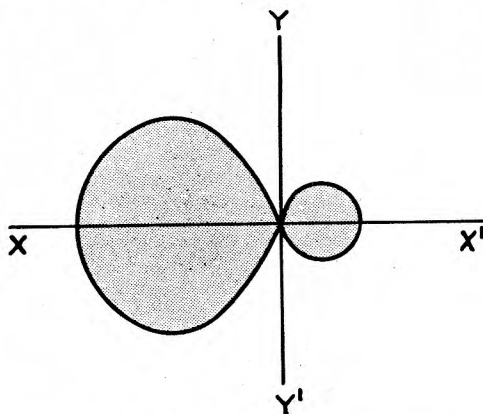


Fig. 14—Axial P.D. of half-wave aerial with reflector

38. It is generally true to say that if the length l is somewhat greater than a half wavelength the relative phases of the currents in the two aerials are such as to cause the total radiation to build up in the direction X in the illustration, and to cancel in the direction X'.

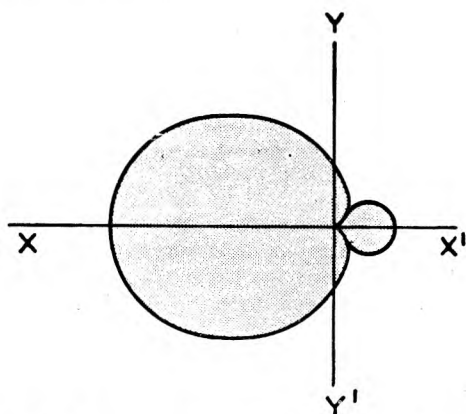


Fig. 15—Transverse P.D. of half-wave aerial
with reflector

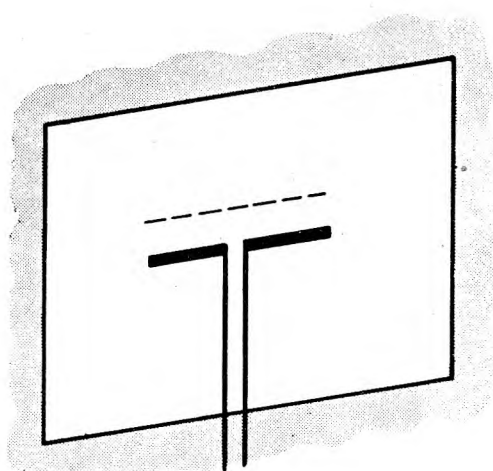
In other words, if l is greater than a half wavelength the parasite acts as a reflector. If l is less than a half wavelength the waves cancel in the direction X and add in the direction X' , so that the parasite acts as a director.

39. Two resultant polar diagrams of an aerial and parasite, when the parasite is acting as a reflector, are shown in fig. 14 and 15. The diagrams show two cross-sections of the polar diagram. Fig. 14 gives the P.D. in a plane containing axis of the aerial, and fig. 15 the P.D. in a plane at right angles to the axis. Some radiation will always find its way back in the direction X' , since the currents in the two aerials will not be of quite the same amplitude and the radiation from the two will, therefore, never quite cancel. The exact shape of the polar diagram will change, of course, as the distance l and d are changed, and it is possible to adjust these distances to give any one of the following conditions:—

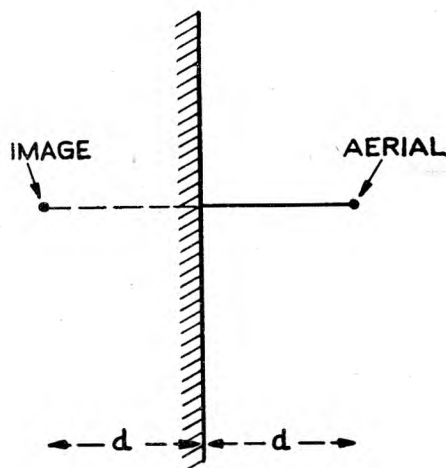
- (1) Maximum radiation in the forward direction (OX).
- (2) Minimum radiation in the backward direction (OX').
- (3) Maximum front-to-back ratio.

These three conditions cannot all be satisfied at one and the same time.

40. One important point to notice is that the beaming is more pronounced in the plane containing the axis of the aerial, see fig. 14, than in the plane at right angles to this, see fig. 15. The best direction of radiation of a system of this sort, that is the direction OX in the illustration, is often referred to as the *line of shoot* of the system.



(a)



(b)

Fig. 16—Aperiodic reflector

41. Parasitic reflectors and directors of the type considered previously are known as *tuned* reflectors and directors. Another type of reflector often used is the *aperiodic reflector*. This is simply a sheet of conductor placed parallel to the aerial as shown in fig. 16 (a). If such a sheet is placed at a distance d from the aerial an image is formed at the same distance behind the sheet, and the radiation at any distant point is obtained by adding the fields from the real and image aerials respectively, taking into account their phase difference.

42. In practice it is usual to employ some form of wire netting instead of a solid sheet of

metal, since it is lighter and has less wind resistance. Very little radiation will escape through the mesh provided that the spacing between the wires is small in comparison with a wavelength. If a single aerial is used with an aperiodic reflector, the forward radiation from the system is maximum when the distance d is a quarter wavelength. It is usual in certain apparatus, however, to use not a single aerial but a loose aerial array with an aperiodic reflector behind. It has been found best in this case to make d either one-eighth wavelength or five-eighths of a wavelength.

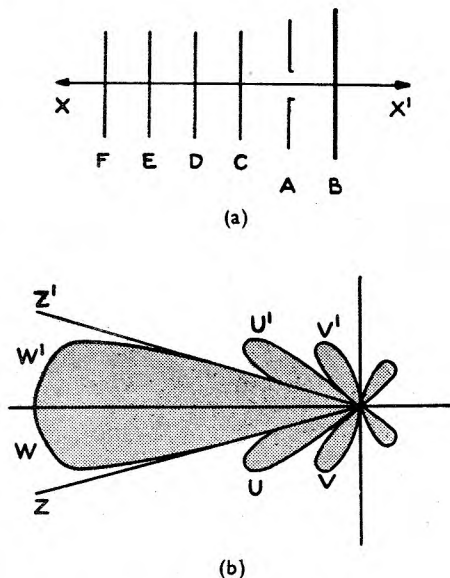


Fig. 17—Yagi aerial

Yagi aeriels

43. A type of aerial known as the *Yagi aerial* is frequently used in radar. It employs a system of tuned directors and reflectors to concentrate the radiated energy into a fairly narrow beam. Fig. 17 (a) shows a typical array of this type, with an energised aerial A, a reflector B, and four directors, C, D, E and F. The aerial A is a driven half-wave aerial fed from a transmitter by a transmission line at its centre. The reflector B is greater in length than a half wavelength, and the directors are shorter than the half-wave aerial. Both the horizontal and vertical sections of the polar diagram of the system are similar to the curve shown in fig. 17 (b). The line of shoot is OX and most of the energy is radiated between the directions OZ and OZ'. In directions further from the line of shoot from OZ and OZ' there are a number of subsidiary maxima in the polar diagram.

44. The principal part of the diagram, OWXW', is known as the *main lobe*, while the subsidiary parts such as those with maxima at U, U', V, V', are known as *side lobes*. A Yagi aerial is particularly useful in aircraft installations working on wavelengths of about 1.5 metres, since it provides a convenient method of obtaining a beam, has very low wind resistance, and is portable.

45. In all these accounts of the action of parasitic aeriels, the systems have been considered as transmitting arrays. From the reciprocity theorem, however, it follows that the polar diagrams shown above will each have its form unchanged if the systems are used for reception.

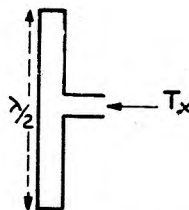


Fig. 18—Folded aerial

Folded aeriels

46. The presence of a large number of parasites in the neighbourhood of a driven half-wave aerial has the effect of lowering its radiation resistance very substantially, and this may cause serious difficulty in matching the aerial into the transmission line which feeds it. The radiation resistance can be increased again, however, by using a folded aerial as the driven element. A folded aerial is shown in fig. 18. It consists of two parallel rods, each a half wavelength long, held close together by two conductors at their ends. The transmission line feeds into an airgap at the centre of one of the rods. The folded aerial has the same polar diagram as an ordinary half-wave aerial, but its radiation resistance is higher. If the two half-wave rods have the same diameter, the radiation resistance is increased by a factor of four, but this factor can be changed by varying their relative diameters, and in practice it can have any value from 2.5 to 7.

Resonant slots

47. For certain purposes it is useful to use not an aerial but a slot for radiation and reception. It can be shown that if a slot such as PQ, see fig. 19 (a) is cut in an infinite conducting sheet, it will radiate in the same way as an aerial. It must be driven by a balanced transmission line feeding into the points RS at its

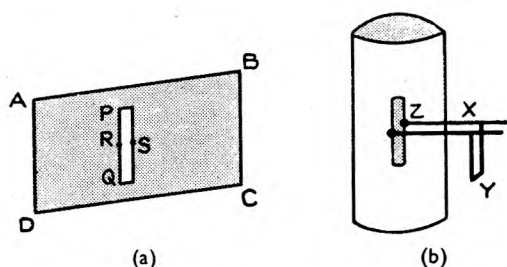


Fig. 19—Resonant slots

centre. It will behave as an impedance at the termination of the transmission line just as an aerial does, and when its length is approximately a half wavelength this impedance will be a pure resistance.

48. A half-wave slot has a polar diagram of exactly the same shape as that of a half-wave aerial. The one difference between the slot and the aerial, however, is that the electric and magnetic fields are interchanged. Thus if the aerial in fig. 10 were to be replaced by a slot, the magnetic field waves would be directed along the line of longitude and the electric field along the line of latitude. It is in this interchange of fields that the value of the slot lies, as we shall see later in dealing with the feeding of apertures.

51. The simple aeriels which have been described previously, including the Yagi aerial, possess polar diagrams that are too broad for many operational requirements; consequently other systems capable of producing a more highly beamed radiation are commonly employed.

52. There is a very general rule which must be followed in designing a highly directional aerial; it may be stated as follows:—

To produce a beam of radiation first distribute the power to be radiated over a large surface and then radiate it from this surface.

The term large surface means a surface whose linear dimensions contain several wavelengths at least.

53. Fig. 20 represents a rectangular surface over which the transmitter power has been spread *uniformly* and from which it is radiated everywhere in the same phase. Let the edges of the surface be a and b as shown. The section

49. It is impossible in practice, of course, to cut the slot in an infinite conducting sheet; but, if the sheet ABCD in fig. 19 (a) is sufficiently large the slot operates very efficiently. A unit such as that shown in fig. 19 (b) is often used. Here the slot is cut in the side of a cylinder, and the cavity behind it acts as a reflector, so that the directional properties are increased.

50. If the slot is required for very wide band work, its bandwidth must be artificially increased. Just as in the case of an aerial, it is possible to do this by making the slot wider. A quarter wavelength of shorted lines can also be used as it was for the aerial, but cannot be connected directly across the slot, since the impedance of the slot varies in the opposite way from that of an aerial when the frequency is changed. Thus as the frequency is increased and the wavelength decreased, the slot becomes too long, but a slot which is slightly too long is capacitive. A shorted stub connected across it would also become capacitive with increase in frequency and so would exaggerate, instead of cancelling, the effect. It would be possible to obtain the required compensation by connecting an open-circuit stub across the slot, but for practical reasons it is usually more convenient to connect a shorted stub XY across the transmission line a quarter wavelength away from the slot. This gives the required result.

BEAMING

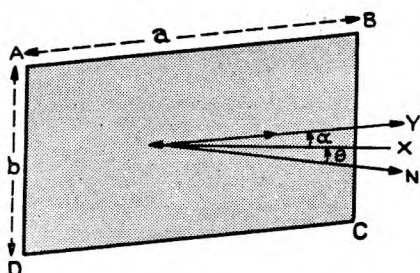


Fig. 20—Uniformly-radiating surface

of the polar diagram by a plane through its axis and parallel to one of the edges (the edge a for instance) is shown in fig. 21. The pattern comprises a main central lobe and side lobes.

54. When the edge a , to which the section of the lobe is parallel, exceeds several wavelengths, that is, $\frac{a}{\lambda}$ is greater than about 4 or 5

the angular distance θ from the maximum of the main lobe to the first zero, on either side is given by the following approximate but useful formula:—

$$\theta \text{ deg.} \approx 60 \frac{\lambda}{a}$$

where the wavelength λ and the edge a are measured in the same units of length. According to the original proviso $\frac{\lambda}{a}$ is small and less than 0.25, but the formula should not be employed when the angle θ exceeds 20 deg.

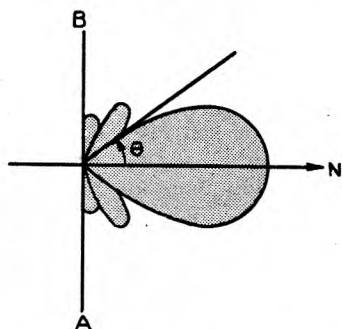


Fig. 21—Axial P.D. for fig. 20

55. The angle θ is a good measure of the beam width. It follows, from the formula that when $\frac{a}{\lambda}$ is large, a high degree of beaming is achieved, as was stated above as a general principle.

56. Instead of using the angle θ from the direction of maximum signal to the first zero on one side, the beam width is often taken to be the angular separation between those two directions, on either side of the main lobe, in which the field strength at a given range is half the maximum on the lobe axis N. This method of measuring beam width leads to a value only slightly greater than the angle θ from N to the first zero. Consequently the approximate formula can be used to cover both definitions where high precision is not required.

57. Fig. 21 has been considered as a section of the radiation pattern parallel to the edge a , but a section parallel to the edge b would have a similar appearance. The angle θ (when $\frac{b}{\lambda}$ is large) is then also found from the formula—

$$\theta \text{ deg.} \approx 60 \frac{\lambda}{b}$$

58. Since the beam widths in the two perpendicular planes parallel to the edges a and b

are determined by the quantities $\frac{a}{\lambda}$ and $\frac{b}{\lambda}$ independently, it follows that the beam may be given any desired form by choosing a and b appropriately.

59. Thus, an aerial for an early warning coastal equipment is constructed to have a large horizontal dimension a and a smaller vertical dimension b . The beam then is narrow in azimuth to give accurate D/F, but is broader in elevation to provide vertical coverage. The beaming in azimuth requires that the aerial be rotated in order that the region in front of the station can be swept to restore coverage in azimuth lost by beaming.

60. For height finding, the radiating rectangle is given a small horizontal dimension a and a long vertical dimension b . The polar diagram is then like a fan, narrow in the vertical plane and broad in the horizontal. By tilting the aerial, the beam can be used to search in elevation.

61. Intense beaming increases the operational range of an equipment, as well as providing accurate measurements of azimuth and elevation. It should be remarked that the approximate formula $\theta \text{ deg.} \approx 60 \frac{\lambda}{a}$, is derived on

the assumption that the power is radiated *uniformly* over the whole surface. If, as commonly happens, more power is supplied to the centre than to the edges of the surface, then the beam is broader than is indicated by the formula. The side lobes are, however, reduced by such a distribution.

62. When power is radiated uniformly from a circular aperture of diameter D then the beam width is given more nearly by—

$$\theta \text{ deg.} \approx 70 \frac{\lambda}{D}$$

63. It can be seen that beaming and ability to rotate become incompatible in an aerial system when the wavelength λ becomes large, for the large dimensions of the aerial system necessary to produce beaming render the system immovable. For instance, the CH system of radar stations which operates on wavelengths of about 11 metres, employs fixed aerial systems with a broad horizontal polar diagram to give coverage in azimuth.

64. By using a vertical stack of end-fed half-wave aeriels and reflectors, moderate beaming is achieved in elevation combined with a broad horizontal polar diagram, which is roughly that of a single half-wave aerial.

65. At the CHL wavelength of $1\frac{1}{2}$ metres, and at lower wavelengths a high degree of beaming is compatible with a manoeuvrable aerial system.

66. According to the reciprocity theorem the polar diagram in reception is also beamed like that in transmission in all cases.

Broadside arrays

67. In this array, which is used in CHL and CGI equipments which both operate on a wavelength of about $1\frac{1}{2}$ metres, beaming is achieved by feeding equal powers in the same phase to individual half-wave aerials which are regularly distributed in a rectangular pattern as shown in fig. 22.

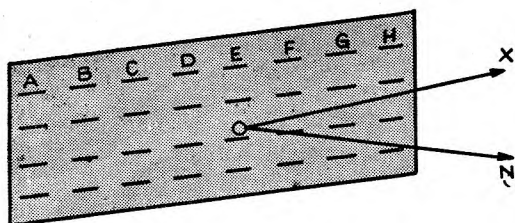


Fig. 22—Broadside array

68. The elementary radiators are arranged in bays, each bay comprising two stacks of 4 radiators. Each pair of half-wave aerials, one in each stack, is end-fed from a feeder line as indicated in fig. 1 (b). Their vertical separation is $\frac{\lambda}{2}$ and the radiators are mounted at a distance $\frac{\lambda}{8}$ from a netting reflector.

69. The CHL array comprises five bays, that is a total of 40 half-wave radiators. The effective breadth of the array is approximately 5λ and the horizontal beam width therefore roughly $\frac{60}{5} = 12$ deg. The correct value is $10\frac{1}{2}$ deg.

Paraboloidal reflectors

70. At higher frequencies it becomes extremely difficult to use a broadside array, since in order to obtain a sufficiently high power gain it is necessary to use a very large number of aerials, and it becomes virtually impossible in practice to match them successfully to a transmission line.

71. To overcome this difficulty, another method of beaming is required, and the most satisfactory type of aperture for wavelengths considerably shorter than 1.5 metres is the parabolic mirror. Such a mirror consists of a reflecting surface of paraboloidal shape, and its action can best be understood by first considering the simple geometrical properties of the parabola.

72. A parabola is the figure traced out by a point which moves in such a way that it always remains equidistant from a fixed line called the *directrix* and fixed point called the *focus*. In fig. 23, ARPB represents a parabola. The line XY is the directrix and the point P is the focus.

73. Any point on the parabola, such as R, is equidistant from both F and XY; that is, the perpendicular from R into XY, namely RT, is equal in length to RF. The line OPF, drawn through the focus and perpendicular to the directrix, is called the *axis*. The point P is called the *pole* of the parabola.

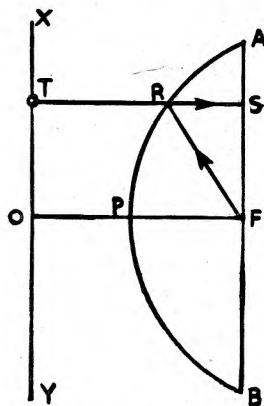


Fig. 23—Construction of a parabola

74. If the parabola is rotated about its axis it will trace out a solid figure which is called a *paraboloid*. When the figure is rotated, the line AFB, which passes through the focus and is perpendicular to the axis, will trace out a circle. This circle is called the *focal plane* of the paraboloid, and the distance PF, which is equal to PO, is called the *focal length* of the paraboloid.

75. Suppose that a source of electromagnetic waves is situated at the focus of a parabolic mirror. Some of the energy will travel directly outwards away from the mirror. The energy which is directed into the mirror

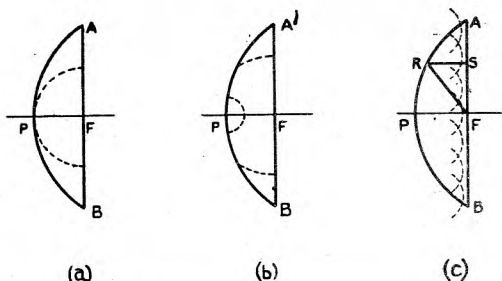


Fig. 24—Reflection from parabolas

will be reflected, however, and after reflection will pass out into space. The way in which this reflection occurs is indicated in the diagram. Consider any line FR drawn from the focus to some point R on the mirror. The line RS is parallel to the axis of the mirror, and is the shortest distance between the point R and the focal plane. It can easily be seen from the figure that the distance:—

$$FR + RS = TR + RS = \text{twice the focal length,}$$

and is, therefore, constant for all positions of R.

76. Consider now a spherical wave crest shown dotted in fig. 24 (a), diverging from the focus and travelling into the mirror. The first part of this wave crest to reach the mirror will be that travelling towards the pole P. As soon as this part of the crest touches P it will give rise to a small ripple, see fig. 24 (b), which will travel outwards from the pole as shown. Meanwhile the original crest will strike the mirror at points farther from P such as R and R', and a series of ripples will diverge from each of those points in turn. A little later the tip of the ripple from P will just have reached the focal plane at F again on its return journey. The total distance it has travelled will then be twice the focal length, that is from F to P and back. Meanwhile the ripples from all such points as R will also have travelled a total distance from F of twice the focal length of the mirror.

77. Thus, from what has been said it will be seen that at the same instant that the ripple from P reaches F, the ripple from R must also reach S, see fig. 24 (c), and the ripples reflected from all other points must arrive at the focal plane simultaneously as shown.

78. These ripples will join to build a plane wave crest across the plane AB, and the reflected radiation at every point in this plane will have the same phase, since the total path (FR + RS) is independent of the original direction FR. In this way the power is distributed over the surface of the aperture and is radiated as a beam

whose widths are given by the approximate formulae $\theta = 60 \frac{\lambda}{a}$ and $70 \frac{\lambda}{D}$ for a rectangular and circular aperture respectively.

79. A complete paraboloid, of course, is infinite in size and, in practice, only a section of it can be used. Suppose that the mirror under consideration is a section of a paraboloid cut across the focal plane AB, so that it includes the portion APB of the complete paraboloid. Such a mirror will obviously act as an aperture, just as did a broadside array. Like a broadside array it will generally have uniform phase distribution across its surface. The same statement would apply, in fact, if the aperture of the mirror did not lie in the focal plane, but in another plane such as MN or RS in fig. 25 (a). It is often more convenient, in fact, not to cut the paraboloid by a plane in this way but to cut a rectangular section from it such as that shown in fig. 25 (b), which, when filled with radiation from source at F, would act as a rectangular aperture.

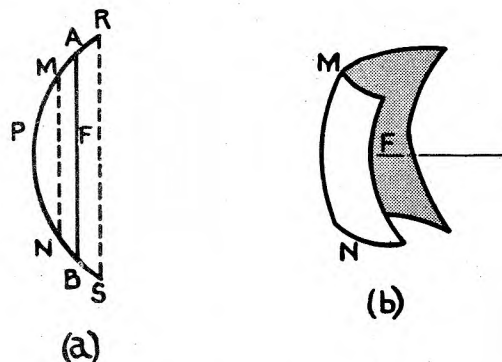


Fig. 25—Parabolic mirrors

Feeding parabolic mirrors

80. Several problems arise in connection with the feeding of parabolic mirrors. Suppose that a mirror of circular aperture is to be fed by a half-wave aerial. The area of the aperture will be determined by the power gain required of the system, but the focal length may be varied at will. Fig. 26 illustrates this fact, showing cross-sections of three mirrors all of the same aperture AB, the first of which has a short focal length, the second a focal length such that the aperture just corresponds to the focal plane, and the third a much longer focal length. The aerial must be placed at the focus, and to avoid the wastage it is necessary to use either a tuned or an aperiodic reflector to throw all the energy into the mirror.

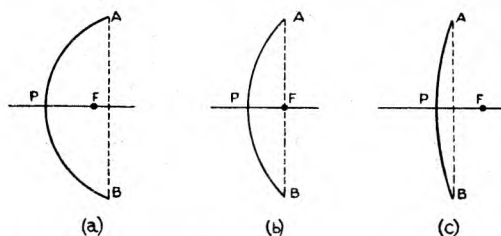


Fig. 26—Optimum focal lengths of mirrors

81. The polar diagram of a single aerial and reflector as shown in fig. 15, whence it is seen that:—

- (1) if the focal length is too small as in fig. 26 (a), the radiation from the aerial will not reach the edges of the aperture, while
- (2) if the focal length is too great as in fig. 26 (c) some of the radiation from the aerial will "spill" over the sides of the mirror and will be lost.

A suitable focal length, in fact, is that shown in fig. 26 (b), and is such that the aperture lies in the focal plane. It is also clear from consideration of the polar diagram of a half-wave aerial and reflector that, with a horizontal aerial, the energy travelling into the mirror will be more highly beamed in a horizontal plane than in a vertical plane, so that the mirror will receive the radiation from the aerial in the way indicated in fig. 27.

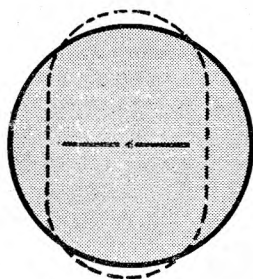


Fig. 27—Feeding circular-aperture mirrors

82. This sketch shows the aperture as it appears on looking into the mirror, while the dotted line is a cross-section of the polar diagram of the aerial in the plane of the aperture. The top and bottom of the mirror will receive more radiation than the sides, so that the feeding will tend to be tapered horizontally. A vertical aerial would, of course, give tapered feeding in the direction AB.

83. This inequality of feeding in the vertical and horizontal directions arises from the fundamental directional properties of a half-wave aerial, and cannot be overcome. Varying the size and spacing of the reflector will, of course, vary the polar diagram of the aerial, and, therefore, the distribution over the aperture. The resultant distribution will, however, never be uniform over the whole area of the aperture, but will always be somewhat tapered in the direction parallel to the axis of the aerial. This example shows that in feeding a mirror there are always two factors to be considered, namely:—

- (1) the polar diagram of the primary source—an aerial with reflector in the above example.
- (2) the amplitude distribution over the aperture of the mirror which is determined by the polar diagram of the primary source, and which in turn determines the final polar diagram of the system.

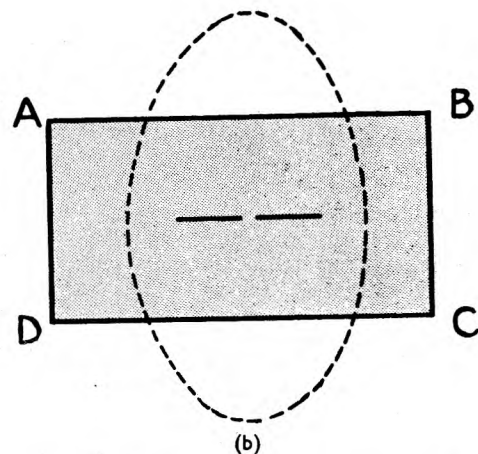
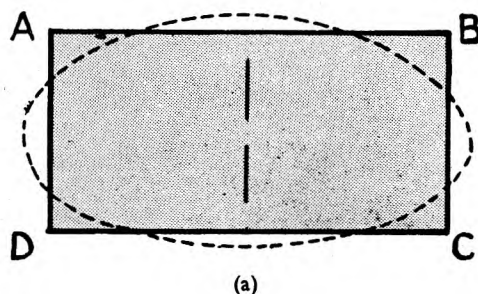


Fig. 28—Feeding rectangular apertures

84. In feeding a rectangular aperture the same difficulty arises. Thus if an aperture such as that shown in fig. 28 (a) is to be fed by a vertical aerial it is relatively easy to fill the whole area without "spilling" too much of the energy over the edges of the mirror, and also without wasting too much of the area in the neighbourhood of the edges by failing to send any radiation in these directions.

85. If a horizontal aerial is used for the same aperture, however, the result is similar to that shown in fig. 28 (b), where a large proportion of the radiation from the primary source misses the mirror altogether, while large areas at the sides go unilluminated. Thus for this type of aperture it is necessary to use a vertical aerial and vertically polarised waves.

86. If horizontal polarisation is to be used it is necessary to employ a resonant slot as the radiating element, since, as previously stated, a slot radiates as an aerial whose electric and magnetic fields are interchanged. If the aperture is too wide horizontally, it may be necessary to divide it into two parabolic sections side by side, each with its own separate source.

87. It will be clear from what has already been said it is not possible in practice to obtain uniform amplitude distribution over the whole of any aperture. This is, if anything, an advantage, since the inevitable tapering reduces the side lobes in the polar diagram.

88. An aperture is often fed by a waveguide. In this case the primary source may be the open end of the guide which has a wide polar diagram, and fills the mirror with radiation in much the same way as an aerial does, since its area is comparatively small, being only of the order of one square wavelength. The wave guide is sometimes terminated in a horn when the aperture of the mirror is longer in one dimension than the other.

89. It is found that, for reasons which need not be discussed in this brief survey, a parabolic mirror functions most efficiently if its focal length is an integral number of half wavelengths. Some difficulty is experienced when an aerial and reflector are used as a primary source, since it is difficult to determine the exact point in the system from which the waves effectively originate, and consequently it is not easy to decide the exact position of the aerial reflector system relative to the mirror.

Beam swinging with parabolic mirrors

90. The direction of the beam from a mirror like that from an array can, of course, be altered by rotating the mirror. This is not always

feasible, however, and the alternative method of altering the phase distribution over the aperture is often used. If the aerial or other source is supported at the focus by means of an arm PF, in fig. 29, and the arm is pivoted downwards through an angle θ as shown, the effect is to advance the phase of the radiation over the lower part of the aperture and to retard that over the upper part, so that the beam is tilted upwards by an amount α . It is found that for small angles of tilt:—

$$\alpha = 0.7\theta \text{ approximately.}$$

It is not possible to tilt the beam by a greater angle than about ± 5 deg. by this method, since the first side lobe becomes troublesome.

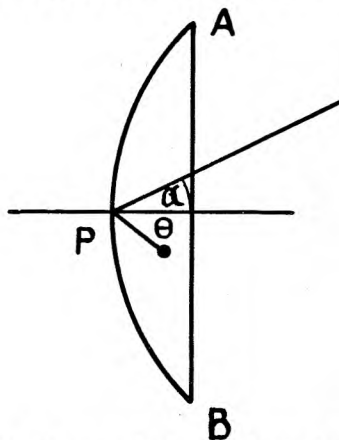


Fig. 29—Beam swinging with paraboloids

Performance diagrams of apertures

91. Consider any aperture used as a common transmitting and receiving unit in radar, and suppose that its polar diagram is similar to that shown in fig. 30. Neglecting the effect of the side lobes, which may be supposed to be small, it would appear, at first sight, that if a target lies in any direction between OZ and OZ', it can be detected by the apparatus, but that if it lies in any other direction it cannot be seen.

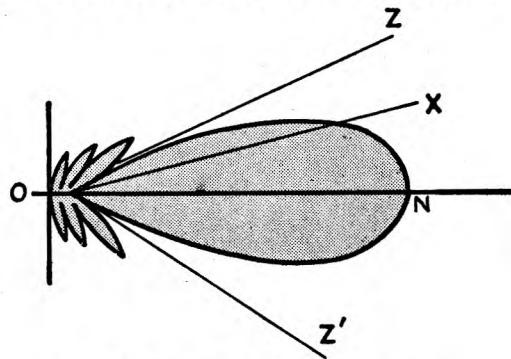


Fig. 30—Theoretical polar diagram

92. In practice, however, when the direction of a target is too far from the line of shoot the return signal may be so weak that it cannot be seen even though it still lies in the main beam; in other words the equipment cannot detect echoes over such a wide range of directions as those included between OZ and OZ', and the effective width of the beam is less than the angle between the two zeroes. It is usually considered, as a working rule, that the effective beam width of an aperture is the width of the beam measured, not from zero to zero, but from half amplitude to half amplitude.

93. In order to see what this means more clearly, consider the aperture first as a receiving system. Suppose that a distant target carries a transmitter working on the same frequency as the radar equipment, and that the aperture receives some of the radiation from this transmitter. If the whole aperture is rotated so that the maximum of its polar diagram passes through the target, the signal received will be greatest when the target lies in the direction ON. When the target is in any other direction OX the received signal will be reduced in amplitude by a factor equal to $\frac{OX}{ON}$.

94. Now, with an actual radar equipment, the target does not carry a transmitter of its own, but scatters radiation which it receives from the equipment. As the beam sweeps round, the energy received and rescattered by the target will vary, and the ratio of the amplitude of waves scattered when the target is in the direction OX to that when the target is in the direction ON will again be $\frac{OX}{ON}$. Thus,

in rotating from ON to OX, the amplitude of the return signal will be reduced on two accounts; first, because the signal received at the target is reduced in the ratio $\frac{OX}{ON}$, and second, because the sensitivity of the receiver has also fallen in this ratio. Thus the resultant return signal will be reduced by a factor

$$\frac{OX}{ON} \times \frac{OX}{ON} \text{ or } \left(\frac{OX}{ON}\right)^2.$$

95. It is, therefore, possible to draw a second type of polar diagram for an equipment which gives the overall change in amplitude with direction, and which is actually the product of the transmitting and receiving polar diagrams. Such a diagram is called a *performance diagram*.

96. In the case of the aperture just considered, half amplitude on the polar diagram corresponds to quarter amplitude of the return signal. In the same way the ratio of amplitude of the side lobes to that of the main lobe is reduced in the performance diagram.

97. The performance diagram has two meanings, similar to those which characterised the polar diagram:—

- (1) it is a measure of the variation of amplitude return signal with direction from a target at constant range, and
- (2) it is a measure of the variation of the square of the range of a target with direction when the target moves in such a way that its signal to noise ratio is constant.

SPECIAL AERIALS

98. There are several types of beamed aerial systems which fall outside the scope of the preceding discussion, and some of these are worth a brief mention.

Cosec θ aerials

99. In certain types of airborne equipment working on centimetre wavelengths, it is necessary to illuminate the ground below with a narrow beam of radiation in the way shown in fig. 31 (a). A narrow strip of ground ABCD extending from the point A beneath the aircraft to a distant point C which may be on the horizon, receives the radiation from an aperture

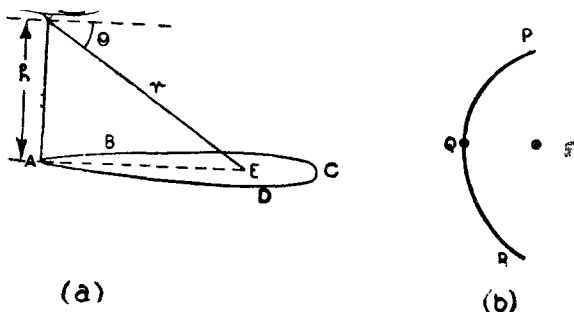


Fig. 31—Cosec θ aerial

housed in a perspex *blister* beneath the belly of the aircraft. Now, the distance from the aircraft to any point E on the ground is obviously $r = h \operatorname{cosec} \theta$, so that to illuminate the ground below uniformly it is necessary to use an aerial whose polar diagram will obey a cosecant law. This is accomplished by an aperture of cross-section similar to that shown in fig. 31 (b). The part AB is circular and has its centre at the

point F, while the part BC is parabolic, and has its focus at F. The aerial or waveguide feed is, of course, situated at the point F, and the whole array usually rotates about a vertical axis through this point, so that the beam sweeps out a circular path on the ground below. In practice it is found that the law $(\operatorname{cosec} \theta)^n$ is preferable where n is greater than 1. The best value for n is not yet agreed upon.

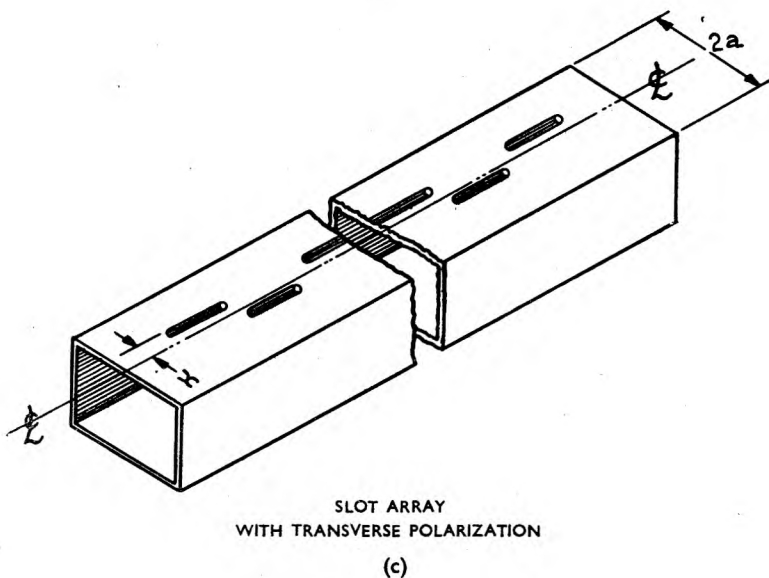
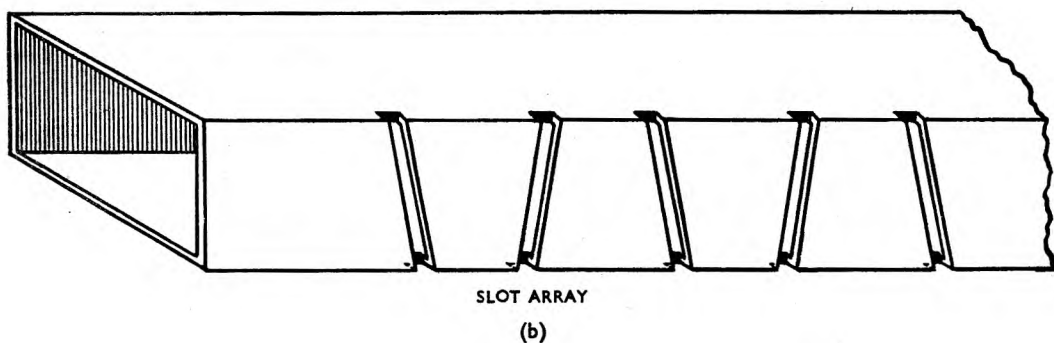
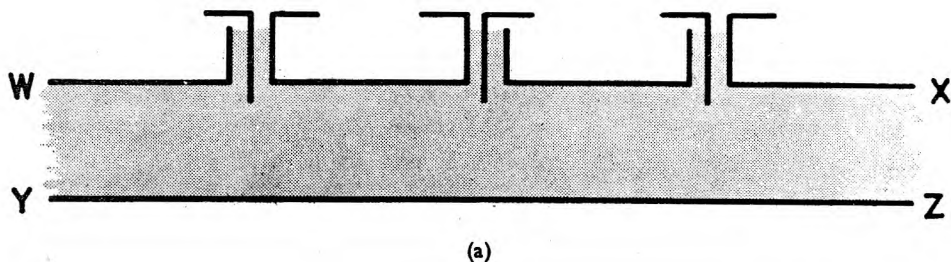


Fig. 32—Waveguide arrays

Polyrod aerials

100. A well-known method of obtaining a narrow beam is to terminate a wave-guide with a rod of polythene or other dielectric material, allowing the rod to protrude some distance beyond the end of the guide. The arrangement acts as a leaky waveguide, and radiates, giving a fairly high degree of beaming, the axis of the beam being in the direction in which the rod points.

Waveguide arrays

101. Fig. 32 (a) illustrates a method of feeding a microwave ($\lambda = 3$ cm. or 9 cm.) array directly from a waveguide. The half-wave aerials are fed from short lengths of coaxial line whose inner conductors extend into the waveguide to form probes parallel to the electric field in the wave. By using a travelling wave and spacing the probes at a separation of $\frac{\lambda g}{2}$, where λg is the wavelength of the wave in the guide, the half-wave aerials or dipoles may be driven in phase. It is also necessary to alternate the sense in which the halves of the dipoles are connected to the inner and outer conductors of the coaxial feeder, as is shown in fig. 32 (a).

102. It is found more convenient, in recent applications, to radiate directly from slots cut in a wall of the waveguide. These are also spaced with a distance $\frac{\lambda g}{2}$ between their centres and are so cut that they intercept the flow of current on the wall. By cutting the slots at the appro-

prate angle across the direction of the wall current, the amount of energy radiated from each slot can be controlled; so that the slots more distant from the source radiate as strongly as the nearer slots although the travelling wave has been weakened by the radiation abstracted from it before it reaches them. Figs. 32 (b) and (c) show two forms of slot array.*

Cheese aerials

103. In the cheese aerial, or parabolic cylinder, the reflecting surface is not a paraboloid of revolution, like the reflectors discussed in para. 63 on, but a parabolic cylinder. It is shown in fig. 33.

104. The reflecting surface is a flat strip which remains straight in one plane at right angles. The "roof" and "floor", shown in the figure, are metal plates. Cylindrical waves are emitted in the focal plane from the end of a waveguide terminated in a horn. These waves are reflected at the back and return to fill the large rectangular aperture. This type of aerial is used with a wavelength of 10 centimetres in a coastal equipment, known as AMES, Type 14, for detecting ships and low flying aircraft.

105. The dimensions of the aperture are 15' $\times$ 2' 6", and when the long dimension is horizontal the beam width in azimuth is only $1\frac{1}{2}$ deg.

106. Such an aerial, with the long dimension in the vertical plane, is used in height-finding equipments.

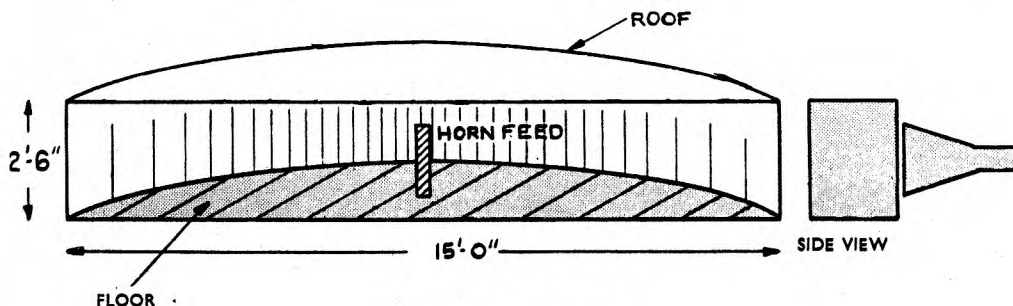


Fig. 33—Cheese aerial

* In practice it is found more convenient not to space the radiators at $\frac{\lambda g}{2}$, but at a slightly greater separation. The beam is then not at right angles to the line of radiators.

EFFECT OF GROUND ON RADIATION PATTERNS

Theoretical aspects

Reflection from conducting surfaces

107. So far as the directional properties of aërials have been considered without reference to the effect of the earth, and the polar diagrams obtained have been *free space* polar diagrams.

108. In practice, however, any radiating or receiving array will be elevated at a certain height above the earth, and its polar diagram will be modified by the effect of ground reflection. This is particularly important in the case of ground equipment, and must now be described in some detail. The problem is complicated by the fact that the earth behaves sometimes as a conductor, and sometimes as a non-conductor in reflecting electromagnetic waves, and the exact way in which the reflection occurs depends on the frequency.

109. Suppose first that the earth behaves as a flat perfectly-conducting surface. Such a surface will totally reflect all electromagnetic waves falling on it. The conditions for reflection will, however, depend on the plane of polarisation of the waves and it will be necessary to take the cases of horizontal and vertical polarisation separately.

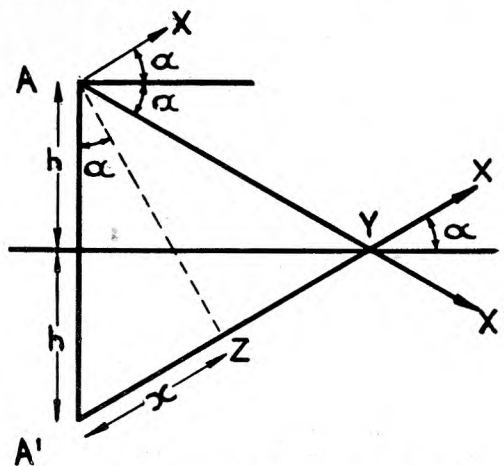


Fig. 34—Reflection from conducting surfaces

110. Suppose that a horizontal half-wave transmitting aerial with no reflectors is situated at the point A in fig. 34; let its height above the earth be h . Consider the illumination from it at a distant point in the direction AX at an angle of elevation α . The axis of the aerial is taken perpendicular to the plane of the paper so that any direction such as AX lying in the plane of the paper will be one of its best directions of radiation, and it will radiate equally well in any

such direction. The distant point will receive two trains of waves, one travelling direct, along the path AX, and one being reflected from the earth and travelling by the path A'YX¹.

111. It is clear that the reflected waves will travel further than the direct waves, and so will generally be out of phase with them. The lines AX, YX¹ in the figure meet at the distant point whose illumination we are considering; and, if this point is sufficiently remote, those two directions will be virtually parallel. The reflected waves will appear to come from an image point A¹, the same distance h below the ground as the aerial is above. It is clear then that the reflected waves will travel a distance x farther than the direct waves before reaching the distant point, and that this distance x will vary as the angle of elevation varies. When α is zero, x will be zero and as α increases x will also increase.

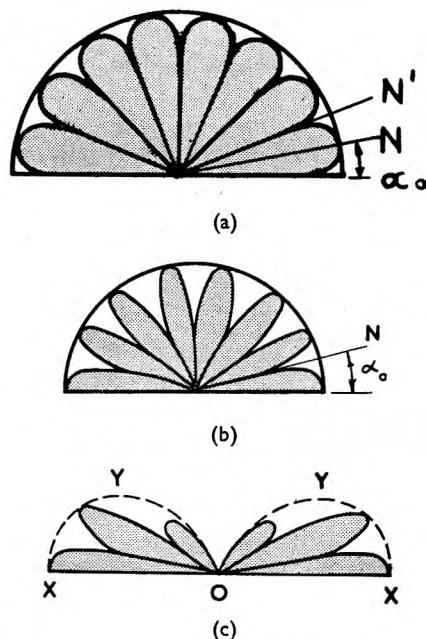


Fig. 35—Lobing due to reflection from conducting surfaces

112. Now, horizontally-polarised waves are reflected from a conductor with phase reversal, so that a trough is reflected as a crest and a crest as a trough. Thus if x is equal in length to one-half wavelength, the reflected waves will have travelled one-half wavelength farther than the direct waves, but will also have effectively lost a half wavelength on reflection. They will therefore, be in phase with the direct waves at

the distant point and the resultant amplitude will be larger. If, on the other hand, x is equal in length to one wavelength the reflected waves will be behind the direct waves by one and a half wavelengths, one wavelength due to extra distance travelled plus a half wavelength on reflection, so that they will be out of phase at the distant point and will cancel.

113. As α and x increase the resultant signal varies, therefore, in the way shown in fig. 35 (a). Since the earth is considered here as a perfect reflector, the direct and reflected waves will be of the same amplitude and will exactly cancel at certain angles of elevation. Along the surface of the ground, where α and x are zero, there will be no path difference between the two waves, but they will be exactly out of phase owing to the reversal on reflection, so that at zero angle of elevation there will be no signal.

114. The number of lobes in the polar diagram depends on the height of the aerial and on the wavelength, and is equal, in fact, to the number of quarter wavelengths contained in the height of the aerial. Thus the shorter the wavelength and the higher the aerial the more lobes there will be in the complete diagram, and the smaller will be the angle of elevation of the first lobe above the ground. The number of lobes is, in practice, usually large, and the angle of elevation of the first maximum α_0 , in fig. 35 (a), is given by the approximate formula:

$$\alpha_0 = 47 \frac{\lambda}{h}$$

where λ is measured in metres, h in feet, and α_0 in degrees. The first minimum ON' , is at an angle of elevation of about twice this value, namely, $94 \frac{\lambda}{h}$.

115. Suppose now that the horizontal aerial at A is replaced by a horizontal slot, which radiates in the same way as the aerial. The waves will now be vertically polarised, and it can be shown that they will again be totally reflected but will not suffer any phase change on reflection. This leads to a polar diagram of the same type as that in fig. 35 (a), but with the maxima and minima interchanged, as shown in fig. 35 (b). Thus the surface of the earth is now illuminated. As before the number of lobes depends on the height and the wavelength, and this time the angle of elevation of the first minimum ON is given by the approximate formula:—

$$\alpha_0 = 47 \frac{\lambda}{h}$$

α_0 being measured in deg., λ in metres and h in feet.

116. A resonant slot has been taken as the source since its action is simpler than that of an aerial. In practice, however, the source of waves will usually be an aerial; and if an aerial is to give vertically polarised waves its axis must, of course, be vertical. If such an aerial is mounted with its centre at a height h above the earth, it will not itself radiate equally at all angles of elevation, but will give maximum signal horizontally and no signal in directions vertically upward and vertically downward. The amplitude of the radiation in the direction OX , fig. 35 (c), will, in fact, vary with α in a manner given by the half-wave aerial polar diagram in fig. 9. Thus the ground reflection lobes will vary in amplitude in the way shown in fig. 35 (c).

117. It can be shown that they are enclosed by an envelope $X'Y'OYX$ and that this envelope is the shape of the polar diagram of a half-wave aerial, in free space. The position of the minima will, however, be unchanged, and the angle of elevation of the first minimum will still be given by the same formula. A vertical slot would have had a similar effect in the case of horizontally-polarised waves.

Reflection from a dielectric

118. For waves of very high frequencies the surface of the earth acts not as a perfect conductor, but almost as a perfect non-conductor of electricity. It is necessary now, therefore, to give a brief account of the reflection of waves from a non-conducting dielectric. Reflection of this type is often met in ordinary optics, well-known examples being the reflection of light from the surface of water or glass. The problem of reflection of electromagnetic waves from a dielectric earth is, in fact, exactly similar to that of light waves reflected from any transparent medium. It will again be necessary to consider the cases of horizontal and vertical polarisation separately.

119. Consider the case of a horizontal aerial at a point A (fig. 34), and suppose now that the reflecting surface is a dielectric whose electrical conductivity is zero. The waves which fall on the reflecting surface will now suffer only partial reflection and will be partly refracted, travelling down into the earth in the direction YX' . The reflection coefficient which is usually given the

symbol ρ , is defined by the formula:—

$$\rho = \frac{\text{amplitude of the reflected wave travelling in the direction } YX'}{\text{amplitude of the incident wave } AY}$$

It depends partly on the dielectric constant and partly on the angle of elevation. At glancing incidence it is always unity, so that at zero angle of elevation the reflection is always total, but at higher angles ρ falls off in the way indicated by the line XY in fig. 36 (a).

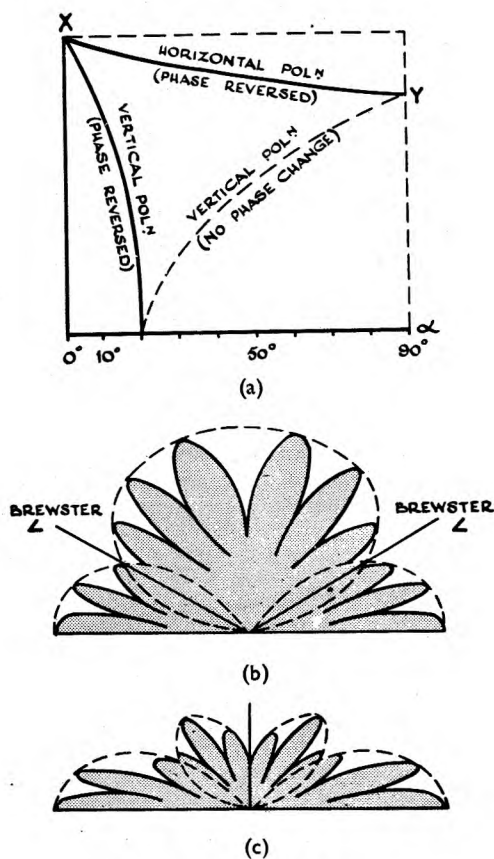


Fig. 36—Brewster angle

120. The slope of this curve becomes less as the dielectric constant increases, and since the dielectric constant of the earth is fairly high, ρ is never very much less than unity. Most types of ground equipment are only expected to "see" aircraft at low angles of elevation, below 30 deg., say; over this part of the curve reflection can be considered to be practically total. The phase of the reflected wave is reversed just as it was for a perfect conductor, and the resultant polar diagram is, therefore, similar to that shown in fig. 35 (a).

121. If the aerial is replaced by a horizontal slot, so that the waves are vertically polarised, the situation becomes more complicated. The reflection coefficient now varies with angle of elevation in the way indicated by the curve XZY in the fig. 36 (a). At low angles of elevation the waves are reflected with phase reversal, but there is an angle known as the *Brewster angle*, at which the reflection coefficient becomes zero.

122. At higher angles of elevation than this, ρ increases again but this time there is reflection with no phase change. It is clear from this that at high angles of elevation the polar diagram will have some similarity to that in fig. 35 (b), while at low angles it will resemble that of fig. 35 (a).

123. One obvious difference arises, however, between this and previous cases. The reflected wave will generally be considerably weaker than the direct wave, so that the two will not completely cancel at the minima. The exact form of the polar diagram is shown in fig. 36 (b). If a vertical aerial is used instead of a slot the whole system is again modified as before, so that the polar diagram is similar to that of fig. 36 (c).

124. The value of the Brewster angle depends on the dielectric constant of the reflecting medium, and is smaller for dielectrics of high constant.

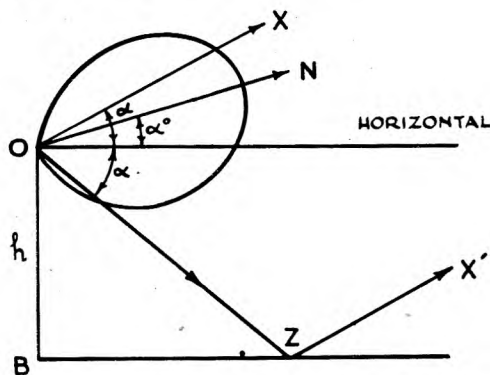


Fig. 37—Ground reflection with beamed radiation

Reflection from the earth

125. The earth is neither a perfect conductor nor a perfect dielectric, so that none of the conditions previously described apply exactly. It is true, however, to a high degree of approximation, to say that for very low frequencies it behaves as a perfect conductor while for very high frequencies it behaves as a perfect dielectric.

126. The transition between these two states depends on the conductivity and the dielectric constant of the material forming the earth's surface, and, therefore, since the sea has a much higher conductivity than has rock or soil, it is different for sea and for land.

127. It is possible to define a frequency known as the *critical frequency* f_c , below which the material forming any part of the earth's surface behaves as a conductor and above which it behaves as a dielectric. This critical frequency is considerably higher for sea than for land, as is shown below.

	Critical frequency (f_c)	Critical wavelength
For sea	900 Mc/s.	33 cms.
For land	1.8 Mc/s.	170 metres

Thus land for reflection purposes invariably behaves as a dielectric at radar frequencies, while the sea, for all except centimetre waves behaves as a conductor.

128. The transition between the two states is not so sharp as may be imagined from the foregoing description. Thus, while at frequencies considerably greater than the critical frequency the earth behaves as a dielectric and at frequencies considerably less than the critical frequency it behaves as a conductor, at frequencies close to f_c the conditions are more complicated and need not be described here in detail.

129. It is true, however, to say that the reflection coefficient is still less than unity and the Brewster angle effect is still observable at frequencies far below f_c . At these low frequencies, the Brewster angle depends not only on the dielectric constant but also on the frequency, and in practice its value can vary very considerably for waves of different lengths, so that it may occur at any angle of elevation between zero and 30 deg. It always occurs at lower angles of elevation, however, for higher frequencies.

Choice of polarisation

130. The plane of polarisation of the waves used in a radar equipment may make a considerable difference to the performance of the equipment, and it is now necessary to discuss briefly some of the more important factors on which the choice of polarisation depends.

131. The reflection of waves from a target depends very largely on the dimensions of the target and the wavelength of the waves, but it is generally true to say that in practice the waves are reflected better if their plane of polarisation corresponds to the direction of the principal lines of the target. Thus, an aircraft or a ship generally presents a silhouette which is long horizontally and short vertically; and, on the whole, these reflect horizontally-polarised waves better than waves which are vertically polarised. Towers and tall buildings on the ground, on the other hand, reflect vertically-polarised waves best.

132. Thus for ground equipments designed to detect aircraft and shipping, horizontal-polarisation will on the whole give better results. In aircraft equipment which is designed to detect buildings and coastlines, however, these advantages of horizontal polarisation are not so obvious.

133. It is found that if vertically-polarised waves are used on such equipments as CHL and AMES Type 11, a large amount of *scatter* is returned from the sea. This scatter is more troublesome during stormy periods. It is considerably less noticeable if horizontal polarisation is used, and for this reason vertically-polarised aerials are seldom employed on these stations.

134. One of the principal problems in all ground equipment is that of *gap filling*. In the polar diagram shown in fig. 35 (a), for example, there are a number of gaps, one at ON^1 , and others at different angles of elevation, and if the transmitting and receiving aerials of a station have polar diagrams of this type they will fail to see a target whenever it disappears into one of these gaps. The usual methods of gap filling will be mentioned later, but it will be clear from what has been said that by using vertical polarisation some of these gaps will be filled or partially filled, as they are in fig. 36 (b) and 36 (c).

135. The most complete gap filling occurs around the region of the Brewster angle, and if this is not at too high an angle of elevation, it may be advantageous to use vertically-polarised waves for this purpose. Most types of ground equipment wish to look at targets at low angles of elevation, however, and the Brewster angle may be too high for useful gap filling. If the earth were a perfect conductor, of course, the use of vertical polarisation would ensure a

good view right down to zero angle of elevation, but in practice this can never be achieved even with sea reflection except at frequencies considerably lower than 100 Mc/s.

136. For certain long-wave equipments the methods used for direction finding and height finding depend on the use of horizontal polarisation, so that in these cases it is not necessary to consider the question.

137. Generally speaking, horizontally-polarised waves are found to be most satisfactory at almost all frequencies for ground equipments. Certain airborne sets work with vertically-polarised waves, and certain special equipments such as Gee and IFF use vertically-polarised waves for gap-filling purposes. Since this section is concerned principally with ground equipment, however, it will be assumed, unless otherwise stated, that any aerial system under discussion is horizontally polarised.

Ground reflection with beamed radiation

138. If an aperture is situated with its centre point at a height h above the ground, the ground reflection lobes will usually appear in the polar diagram. The exact nature of the effect, however, depends on the tilt of the beam.

139. Suppose, first, that the beam is tilted upwards at an angle of elevation α as shown in fig. 38 (a), and suppose that α is larger than the half beam width. In this case none of the radiation from the aperture will reach the ground so that there will be no complications due to reflection.

140. If the aperture is tilted at some smaller angle than this there will be some waves reflected, but, unlike the case of a single aerial above the ground, the directed and reflected waves will not have equal amplitudes. This is clear from fig. 37 which represents an aperture whose best direction of radiation ON is tilted at an angle α_0 . The aperture is at a height h above a reflecting surface BZ, and the radiation at any angle of elevation will include the direct wave OX and the reflected wave OZX¹. The closed curve represents the free space polar diagram of the aperture. The direct wave will obviously have an amplitude OX while the reflected wave will have an amplitude OY. These amplitudes are not equal, so that we will again have the partial interference effect seen in fig. 36 (b) and 36 (c).

141. As the angle α becomes smaller, however, the amplitudes of the direct and reflected waves will clearly become more nearly equal, so that the maxima and minima of the ground reflection lobes will become more

pronounced, until at zero angle of elevation the two will be of equal amplitudes, ON. For high values of α , OY will become first negligible, and then actually zero, so that no ground reflection effect will be seen. The resultant polar diagram is similar to that shown in fig. 38 (b).

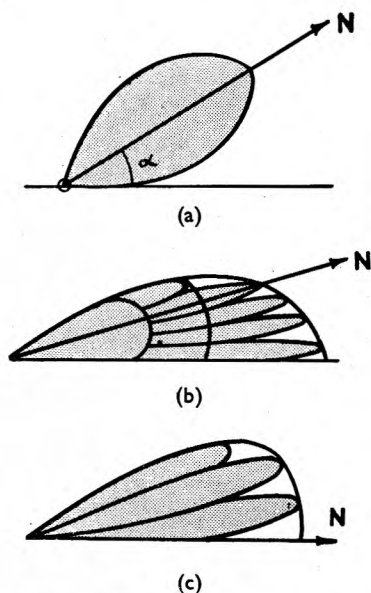


Fig. 38—Lobing due to ground reflection

142. As α_0 becomes smaller, or as the beam is tilted down, the ground reflection becomes more pronounced, until when α_0 is zero and ON is horizontal, it is clear from the symmetry of the beam that the direct and reflected waves will have equal amplitudes at all angles of elevation, and will give more to complete ground reflection lobes as shown in fig. 38 (c). It is, however, clear that the lobes will not all be of equal size but will be enclosed in an envelope whose shape is that of the free space polar diagram of the aperture. This happens, of course, for exactly the same reason as that described in connection with the vertical half-wave aerial above.

143. Thus as α_0 decreases from a large value the ground reflection lobes gradually appear, at first only partially and at low angles of elevation, and, later, completely. The point at which they first appear is theoretically that at which α_0 is equal to the half-beam width, but in practice it is found that they do not become important until α_0 is reduced to about half of the half-amplitude beam-width.

144. When these ground reflection lobes occur the power-gain of the array is increased, since the total radiation along the axis of each

lobe includes both direct and reflected waves. When α_0 is zero and the lobes are complete, the power-gain along the axis of the lowest lobe, which is also the largest lobe, is increased by a factor of four. For partial lobes when α_0 is not zero the increase in gain is less than this.

Additional phenomena of propagation

145. The effect of ionospheric reflection, and the part it plays in ordinary broadcasting in the propagation of electromagnetic waves round the curved surface of the earth, is well-known. It will appear at first sight that the ionosphere would enable radar waves to pass beyond the optical horizon, but this is not generally true, since waves of shorter wavelength than about 10 metres are not usually reflected by the ionosphere under normal conditions. At CH stations, however, the ionospheric reflection leads to a phenomenon known as *long range scatter* which consists of a mass of permanent echoes returned from distant points on the earth's surface by waves which have been reflected by the ionosphere.

146. This type of scatter usually occurs at ranges between 2,000 and 4,000 miles, the precise range depending on the atmospheric conditions, and increasing at times to 8,000 miles. For this reason it is necessary to use a very low pulse-recurrence frequency on CH stations, so that scatter will be returned during the blackout period and will not be seen on the following trace.

147. The usual pulse-recurrence frequency employed is 25 cycles per second, so that the time between consecutive pulses is 40 milliseconds. The trace takes two milliseconds to complete its journey across the tube, so that the blackout time is 38 milliseconds, and is sufficiently long to allow for the return of any long-range scatter up to a range of 4,000 miles. Under phenomenal conditions, when the scatter is returned from longer ranges than this, it is possible to reduce the pulse-recurrence frequency to 12.5 cycles per second.

148. Since the minimum range of the scatter is always of the order of 2,000 miles, it is possible to have a number of stations working on the same radio frequency provided that they transmit in rotation, the second station transmitting immediately the trace of the first has completed its journey, the third transmitting immediately the second trace has completed its journey, and so on. In this way it is possible to

fit in some eight or ten stations so that the blackout period of the last has commenced before the long-range scatter begins to return from the first. All the scatter is then received before the first station retransmits.

149. Stations working on shorter wavelengths than CH are not troubled by long-range scatter, although they may experience other phenomena included under the general heading of scatter. Spurious echoes often appear, for example, and may remain for some considerable time amounting to hours, or may disappear after a few seconds.

150. Any local discontinuity in the dielectric properties of the atmosphere may give an echo. Thus a meteorite entering the earth's atmosphere ionises the air as it passes downwards, and this ionisation may persist for some time and be seen by a radar station. Scatter due to a meteorite usually appears as a well-defined echo at a range of about 80 miles, which disappears after some ten or twenty seconds. Echoes have been received from ionised clouds, and even from large flocks of migrating birds.

151. Although the waves used in radar cannot usually travel round the earth's surface by reflection from the ionosphere, it is often possible to receive echoes from objects below the optical horizon. This happens for a number of reasons. First, if one takes into account the curvature of the earth in calculating the positions of the maxima and minima of the ground reflection lobes, it can easily be shown that the lobes tend to "droop", or bend downwards. This drooping is more pronounced in the lower lobes, as, if the first lobe is at a very low angle of elevation, it may bend round beyond the optical horizon.

152. A second reason for the drooping of the lobes is the refraction of the waves in passing through the atmosphere. The upper layers of air are less dense than the lower strata, and also contain less water vapour; consequently the waves travelling through are propagated in a curved path rather than rectilinearly. This effect may be very marked with certain vertical distributions of temperature and water vapour, and large prominent echoes may be seen from hills many miles away.

153. At very short wavelengths the variation in refractive index of the air over the surface of the sea may cause the first few hundred feet of air to act rather like a waveguide. The waves

after reflection from the surface of the sea are bent round to be re-reflected again and again, so that surface vessels can be seen far beyond optical range.

Limitations to performance of ground equipments

154. It has been seen that with any ground equipment there will always be gaps in the vertical polar diagrams of the transmitting and receiving aerials due to the effect of ground reflection. The lowest of these gaps is at zero angle of elevation, which means that any target at ground level or a short distance above the earth's surface will not be seen by the equipment. The angle of elevation of the lowest lobe can, however, be reduced indefinitely by either raising the aerial or reducing the wavelength, so in practice it is possible to see down to almost indefinitely low angles, and certainly to detect very low-flying aircraft and surface craft on the sea. The higher gaps in the polar diagram are more troublesome, however, and if complete coverage is to be obtained, they must be filled in some way.

155. One method of gap-filling has already been mentioned. If vertically polarised arrays are used, the higher gaps are partially filled, but this is only useful for certain frequencies and heights of aerial, and has for many purposes a number of serious disadvantages.

156. The usual method of gap-filling is to use a second aerial at a different height above the ground. If the heights of the main aerial array and of the second array are correctly chosen, the maxima of the ground reflection lobes of the lower array will correspond with the minima of the lobes of the higher array. Thus by switching from one array to the other it is possible to see a target at all angles of elevation. This method does not, of course, fill the gap at the earth's surface, and there is still a minimum angle of elevation below which the target cannot be seen.

157. Another method of filling the higher gaps is to use two apertures or separate arrays of aerials one above the other, but instead of switching from one to the other in the way indicated above, to reverse the phase of either the upper or lower array. The combined polar diagram of the two arrays when they are fed in antiphase has maxima which corresponds with the minima of the combination when the two apertures are fed in phase. Thus by switching from phase to antiphase the gaps can be filled. In this case, again, the lowest gap still remains.

158. With certain beamed arrays used for shorter wavelengths, partial gap-filling can be

obtained by tilting the beam at a small angle α_0 to the horizontal. The polar diagram will then be similar to that shown in fig. 38 (b), so that there will be no angle of elevation for which the amplitude is zero.

159. It is possible, of course, to fill gaps completely, using a beamed array, by tilting the beam sufficiently to reduce the ground reflection lobes to zero as in fig. 38 (a). In this case, however, the low cover will be lost since the amplitude at low angles of elevation will be very small.

160. The low cover can only be obtained by this method if the beam is sufficiently narrow, but then the high cover will be lost, and to obtain complete coverage with a very narrow beam it must be swept continuously up and down over the range of angles of elevation required. With a wider beam the permanent tilt required for partial gap-filling must not be so great, that the lowest ground reflection lobe is reduced in size; otherwise the amplitude at low angles of elevation will be too small.

161. Gaps due to ground reflection appear both in the polar diagram of the transmitting aerial and in that of the receiving aerial. Thus in order to obtain relatively gapless coverage it is necessary to mount both of these aerials at the same height above the ground, since if the two are at different heights their polar diagrams will not be similar and the final performance diagram will be very "gappy" indeed. In most of the later types of ground equipment this difficulty is automatically overcome since common transmitting and receiving aerials are used in any case.

162. In siting a ground radar station it is necessary to take into account not only the elevation of the aerials, but also any peculiarities of the surrounding country. Thus it is often possible to achieve good gap-filling by siting a station on the crest of a long slope down to the sea, so that at low angles of elevation there are two reflected waves, one being reflected from the slope and one from the surface of the sea. The probability of these two waves and the direct wave interfering in such a way as to cancel completely is very small, and at all low angles of elevation there will be some resultant signal.

163. Again, if a station is to be sited on the top of a cliff, it is possible to find an optimum distance from the cliff edge at which the aerials should be situated so that the best use can be made of the diffraction or bending of the waves as they pass over the cliff edge.

DIRECTION FINDING

Ground equipments

164. In order to define the position of a target on a map it is necessary to know not only its range but also its bearing from the station. Ground radar equipments are, therefore, usually able to measure bearing or azimuth as well as range.

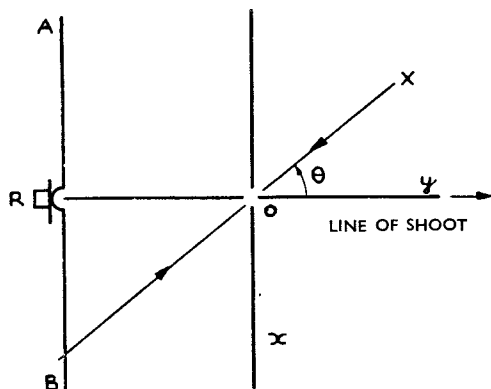


Fig. 39—Crossed dipoles

165. The earliest method of measuring azimuth (θ) is still used in the CH type of station. The transmitting array of such a station is not very highly directional, but floodlights a wide area in front of the station, so that even if a target is a long way off the line-of-shoot it will still receive a sufficiently strong signal to return a good echo to the station.

166. The receiving system consists essentially of a pair of horizontal centre-fed half-wave aerials shown in plan in fig. 39. These aerials are set at right angles, one, the x aerial, having its axis perpendicular to the line of shoot of the transmitter, and the other, the y aerial, having its axis parallel to the line of shoot. If the target lies in the direction OX from the station, the returning waves will set up an alternating voltage in both the x aerial and the y aerial. The amplitudes of the two voltages will not be equal, however, but each will depend on the angle θ between the direction OX and the line of shoot. The way in which each varies with θ will, of course, be given by the half-wave aerial polar diagram.

167. It is possible to find the angle θ by measuring the ratio of these amplitudes. A *goniometer* is employed to measure the ratio.

This instrument consists essentially of two coils set at right angles, the one being connected by a transmission line to the x aerial and the other in the same way to the y aerial. Thus in the *goniometer* there will be two high-frequency alternating fields at right angles, one directed along the axis of the x coil and proportional in amplitude to the voltage set up in the x aerial, the other directed along the axis of the y coil and proportional to the voltage in the y aerial.

168. A third coil, known as a search coil, set between the other two coils, can be freely rotated. A voltage will be induced in this coil owing to the fluctuating magnetic field, and as the coil is rotated this voltage will vary. The signal picked up in this way is fed to the receiver. The search coil is rotated until no voltage is developed in it; that is, until the signal entirely disappears from the display tube. The ratio of the voltages of the aerials can be ascertained from its position.

169. There is an ambiguity in this method of direction finding, since if the signal arrives from the direction shown as YO in the diagram this direction differing by 180 deg. from the direction XO, the ratio of the signals will still be the same. In order to resolve this ambiguity a reflector is placed behind the crossed aerials. This reflector consists of two lengths of conductor joined in the centre through a relay R which is controlled by a press button on the receiver. This relay is normally open and the reflector is not operative. To discover whether an echo is in front of the station or behind it, the operator closes the relay. If the response is from an aircraft in front of the station the amplitude of the echo will then increase, while if it is from an aircraft behind the station the amplitude will decrease.

170. In this method of measuring azimuth a number of serious inaccuracies arise. These are partly due to unavoidable faults in the apparatus, the most serious of which is probably attenuation in the feeder system; and partly due to hills and other peculiarities of the land in the neighbourhood of the station which change the direction from which the waves reflected from the ground approach the aerials. These errors vary from one azimuth to another, and are very difficult to estimate theoretically. In practice it is necessary to find the error at different

azimuths experimentally by flying an aircraft in known directions from the station and finding the amount by which the goniometer reading differs from the true bearing in each case.

171. With more highly beamed arrays, it is possible to measure the azimuth of a target with a much higher degree of accuracy than that attained by the goniometer method. Unless the target lies almost in the beam of the station it cannot be seen, and as the aerial array is rotated, the echo will appear at its maximum amplitude when the aerials are looking directly at the target. The narrower the beam, of course, the more accurate is this method of direction finding.

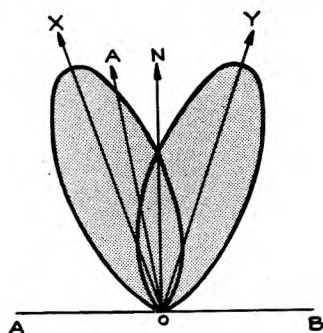


Fig. 40—Split-beam method

172. It is possible to improve the accuracy still further. Suppose that the phase distribution over the aperture from which the beam originates is such as to deflect the beam slightly to one side, so that if AB in fig. 40 represents the plane of the aperture, the line-of-sight lies not in the direction normal to the aperture, but in the direction OX. A target whose bearing is OA will return some radiation to the receiver since its direction lies in the main lobe of the polar diagram. If the phase distribution over the aperture is now reversed so that the line-of-sight corresponds to the direction OY, when the angle between OY and ON is equal to the angle between XO and ON, the echo from A will now have a smaller amplitude.

173. By changing the phase distribution rapidly so that the line-of-sight changes for every alternate pulse, lying in the direction OX for the first pulse, in the direction OY for the second, in the direction OX for the third, and so on, it is possible to estimate the angle between OA and ON by comparing the amplitudes of alternate echoes. If the array is turned until two echoes are equal in amplitude, the normal to the aperture ON must correspond to the direction of the target OA.

174. The advantage of this *split-beam* method lies in the fact that it will measure azimuth very accurately without using a very narrow beam. It is used in certain equipments such as GL, which need to employ fairly wide beams in order to locate targets easily. Many equipments, such as CHL, AMES Type 11, and CHEL use narrow beams and can obtain sufficient accuracy without the use of split technique.

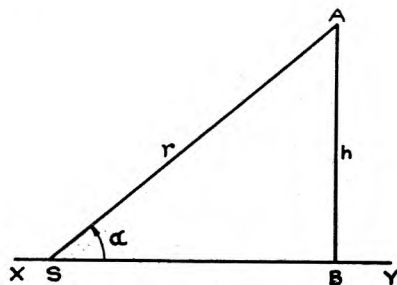


Fig. 41—Elevation finding—short ranges

Airborne equipments

175. The earliest forms of airborne radar equipments requiring D/F were AI Mk. IV and ASV Mk. II. These equipments both radiate forwards in a broad polar diagram, but carry a receiving aerial on each wing. The receiving aerials have crude directional properties, and by pointing the axes of their horizontal polar diagrams slightly outwards to left and right respectively, it is arranged that the response from the right-hand aerial exceeds that from the left when the target is to the right of the line of flight, but is less when the target is to the left. When both aerials give equal responses the target is directly in front on the line of flight. The responses are switched in turn to the cathode ray tube and their sizes directly compared.

176. Airborne equipments working on wavelengths of several centimetres employ paraboloidal reflectors, for common T/R. At these wavelengths such reflectors produce a highly-beamed radiation, and consequently, directions can be measured with precision from the orientation of the axis of the beam.

177. The ability to produce a high degree of beaming, without radiation in unwanted directions, renders centimetre wavelengths highly suitable for use in airborne radar equipments.

ELEVATION FINDING

178. Certain types of ground equipment are capable of measuring the angle of elevation (α) of a target and hence of finding its height.

179. In fig. 41, XY represents a portion of the earth's surface, which for the moment is considered to be flat, and S is a radar station. The angle of elevation α , of an aircraft at A, distance r from S, and at a height h above the ground is given by the formula—

$$h = r \sin \alpha.$$

180. Since h is usually measured in feet and r in miles this formula is more usually written—

$$h = 5280 r \sin \alpha$$

5280 being the number of feet in one mile. If the angle of elevation is small this reduces to the approximate formula—

$$h = 92.5 r \alpha$$

where h is measured in feet, r in miles, and α in degrees.

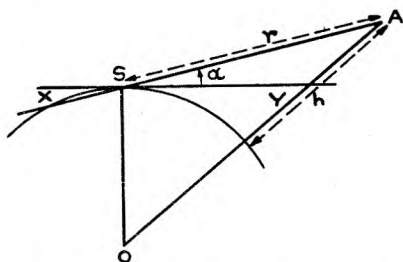


Fig. 42—Elevation finding—long ranges

181. For aircraft at short ranges this value for the height, obtained on the assumption that the earth's surface is flat, is sufficiently accurate. At ranges of more than about 20 miles, however, the effect of earth curvature must be taken into account. Fig. 42 shows a station S on the curved surface of the earth. The line XS is a tangent to the earth at the point S, and in this diagram the height of the aircraft at A, given by the line AB, will be greater than that obtained in the previous case. It can be shown now that

$$h = 5280 \left(r \sin \alpha + \frac{r^2}{7920} \right)$$

where h is again the height in feet, r the range in miles and 7920 is the diameter of the earth in miles. For low angles of elevation this formula becomes approximately—

$$h = 92.5 r \alpha + \frac{2}{3} r^2.$$

A.P.2897K, Height-Range-Elevation Tables, gives values of α for various values of r and h evaluated for radar purposes.

182. It is clear, then, that the height of an aircraft can be measured if its angle of elevation and range are found. The method of measuring range is well known and a brief description of methods of finding the angle of elevation will be given in the following paragraphs. It is necessary to add that the effect of atmospheric refraction has to be considered as well as the effect of the earth's curvature and the above formulae must be modified slightly on that account.

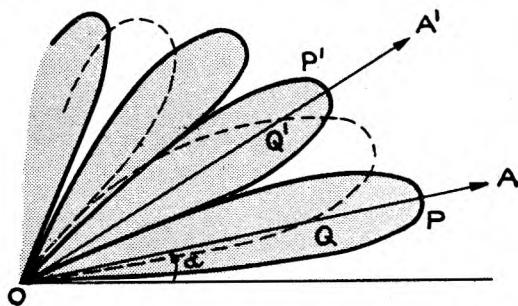


Fig. 43—Classical method of elevation finding

Classical method

183. The original method used in radar for measuring elevation was based on the comparison of signals received in aerials at different heights above the ground. Consider a receiving aerial at a height h , above the ground. Its vertical polar diagram will be similar to that shown by the full line in fig. 43 (a), and will consist of a number of ground reflections. Thus an aircraft in the direction OA, at an angle of elevation α , will return an echo whose amplitude is proportional to the length of the line OP.

184. A second lower aerial, whose height is h_2 feet, will have a vertical polar diagram similar to that shown by the dotted curve, the lobes being wider and fewer in number. The aircraft will set up in the lower aerial a signal whose amplitude is proportional to the length of the line OQ.

185. The ratio of the signals in the two aerials, $\frac{OP}{OQ}$, will vary as the angle α varies and to every angle of elevation there will correspond a definite ratio. By comparing the two signals it is, therefore, possible to measure α .

186. This method of height finding is used in GL, CH, CHB and GCI. In CH and GL the goniometer is used to compare the signals. In CHB and GCI they are compared visually, being displayed side-by-side on a cathode ray tube.

Disadvantages and limitations

187. The classical method is subject to ambiguities, since there are a number of angles of elevation corresponding to any ratio. Thus if the aircraft is in the direction OA for instance, the ratio $\frac{OP'}{OQ'}$ is equal to $\frac{OP}{OQ}$. This particular ambiguity can be resolved, since it can be shown that if the aircraft is in the direction OA which lies in the lowest lobe of both aeriels, the signals in the two aeriels will be in phase, whereas if it is in the direction OA' which lies in the first lobe of the lower aerial and the second lobe of the higher aerial the signals will be exactly in antiphase. It is possible both by the goniometer method and the method of visual comparison to compare the phases of the two signals as well as their amplitudes, so that it is easy to discover which is the correct angle of elevation.

188. At higher elevations, however, further ambiguities occur and there may be a number of angles at which the ratio of OP to OQ has the same value again. This is not very serious, since it is usually possible to arrange that the higher ambiguities are at such a large angle of elevation that when they occur the aircraft will have to be either so close to the station that it will be lost in the permanent echoes or at an impossibly great height.

189. Distortion of the polar diagram caused if the ground near to a station is not quite level, is a more serious disadvantage. This is particularly troublesome on CH stations which look down to low angles of elevation so that the ground reflection may occur at distances as great as one or two miles from the station and at which it is almost impossible to obtain a perfectly flat surface of sufficient extent to give a normal polar diagram. It is not so serious in GCI work since the stations have low aeriels and do not look down to low angles, so that ground reflection is limited to a few hundred yards. In GL stations, which have to look to very high angles of elevation, the reflection area is so limited that it is possible to use an artificial earth consisting of a horizontal sheet of chicken netting, surrounding the aeriels at a small distance above the ground.

190. All equipments which use the comparison method of height finding, however, must

in practice be calibrated, since it is impossible to calculate the polar diagram sufficiently accurately. To calibrate a station it is necessary to obtain a graph of ratio against angle of elevation either by performing a test flight with an aircraft or by flying a balloon at various heights.

Height estimation from performance diagram

191. It was shown that the performance diagram of an equipment could be regarded as a measure of the variation of the square of the range of a target with direction when the target moved in such a way as to keep its signal-to-noise ratio constant.

192. Suppose that an aircraft moves in space in such a way that it remains always at maximum range from the equipment, in other words, so that the signal-to-noise ratio is always unity. It is possible to draw a performance diagram which will show the relationship between the square of the range and the direction of the aircraft, i.e., a graph of r^2 against direction. From this diagram it is also possible to obtain a second graph which will give the relationship between $r_{\max}$ and direction.

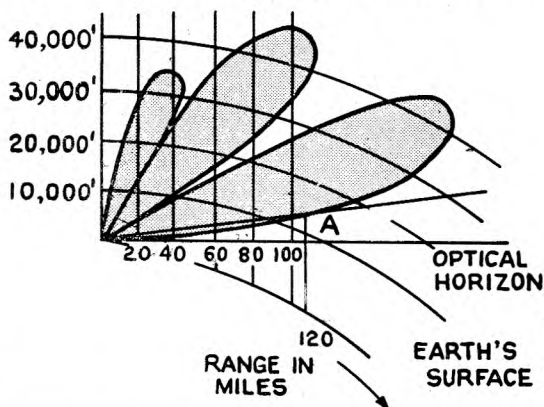


Fig. 44—Height estimation from performance diagram

193. Such a graph is shown in fig. 44. Ranges are marked along the curved surface of the earth and heights are shown in thousands of feet. The height of an aircraft can be estimated approximately by using this diagram from the range at which it is first seen as it flies in towards the station. Suppose for example, it is first picked up at a range of 110 miles as shown. This means that the signal-to-noise ratio is unity at this range, and from the diagram it is clear that a signal-to-noise ratio of unity at 110 miles corresponds to a height of 20,000 feet. Performance diagrams such as this are used on CH and GCI stations for estimating heights.

Modern methods

194. As mentioned above, the classical or ground reflection method of height finding is unsatisfactory and the need for a more precise and reliable method was appreciated from the first.

195. As the limitations of the classical method are attributable to its use of the ray reflected from the ground, the obvious remedy is to employ an aerial system that produces a high degree of beaming in the vertical plane. The aerial must also be able to tilt to any required angle so that the beam can search in elevation to restore the vertical coverage lost by the beaming. The angle of elevation is then that of the beam axis at which the aircraft response is greatest.

196. The realisation of the method awaited the development of radar on the shorter wavelengths of 50 centimetres and 10 centimetres, since the vertical dimensions required for aerials on the longer wavelengths are prohibitively large.

197. The first equipments using this principle were actually operated on a wavelength of $1\frac{1}{2}$ metres. They comprised a vertical stack of 56 dipoles mounted along a CH wooden 200 ft. receiver tower. This equipment was called VEB (Variable Elevation Beam), and was very cumbersome.

198. The next height-finding equipment (DMH) operated on a wavelength of 50 cm. It employed a large netting paraboloidal reflector 30 feet long vertically.

199. The most recent and very satisfactory equipments, classed as CMH (Centimetre Height) operate on a wavelength of 10 centimetres and employ a cheese aerial with its long dimension (15 ft.) in the vertical plane. This aerial produces a fan-like beam $1\frac{1}{2}$ deg. in breadth vertically and 10 deg. broad in azimuth. The mirror is tilted to scan in elevation and can be rotated to search on any azimuth. These equipments are more fully described in A.P.1093C, Chapter 4, and A.P.2525W.

POWER GAIN OF AERIAL SYSTEMS

Mathematical aspects of polar diagrams

200. Suppose that an aerial system is placed at O (fig. 8) and that the radiation is examined over the surface of a large sphere with O as centre. In general there is a direction ON in which the field possesses its greatest strength. Let the field strength at N be E_0 .

201. The field strength E at another point P on the sphere at the same distance r from O as N, depends only on the latitude θ and the longitude φ of P, and is given by the equation—

$$E = E_0 f(\theta, \varphi)$$

where the function $f(\theta, \varphi)$ of the latitude θ and longitude φ is the mathematical representation of the polar diagram of the radiator. Clearly, $f(\theta, \varphi)$ equals unity when θ and φ are each equal to zero in fig. 8, but is elsewhere less than or equal to unity.

Power carried by electromagnetic waves

202. The electromagnetic wave travelling outwards across the spherical surface carries power away from the radiator O. To estimate this power, a theorem of electromagnetism called Poynting's theorem is employed.

203. According to this theorem the average flux of power across an area S square metres perpendicular to the direction of propagation

of the wave whose field amplitude is E volts per metre is—

$$\frac{E^2}{240\pi} \text{ S watts.}$$

Thus the flux across an element of area dA sq. metres at P is—

$$\frac{E^2 dA}{240\pi} \text{ watts,}$$

and the flux of power W over the whole spherical surface is—

$$\begin{aligned} W &= \int \frac{E^2 dA}{240\pi} = \frac{E_0^2}{240\pi} \int f^2(\theta, \varphi) dA \\ &= \frac{E_0^2 r^2}{240\pi} \int_{-\frac{\pi}{2}}^{\frac{\pi}{2}} \int_0^{2\pi} f^2(\theta, \varphi) \cos \theta d\theta d\varphi \text{ watts.} \\ &\quad (\text{since } dA = r^2 \cos \theta d\theta d\varphi). \end{aligned}$$

204. Suppose the polar diagram to be unbeamed so that equal fluxes of power $\frac{E_0^2 dA}{240\pi}$ stream across all elements of area dA irrespective of position in the surface of the sphere. Then the total power passing across the sphere becomes—

$$W_0 = \frac{E_0^2}{240\pi} \times 4\pi r^2 \text{ watts.}$$

A source radiating *isotropically* in this fashion is called an *Isotropic Radiator*. It is a fictitious source unless the radiation is everywhere unpolarised.

205. To produce a field E_0 at range r equal to the maximum field E_0 of the actual radiator at the same range, the isotropic radiator requires power W_0 , and the actual radiator power W . The ratio W_0/W is defined as the *power gain*, G , of the actual radiator, with respect to the isotropic radiator as standard. The isotropic radiator is wasteful of power if maximum field strength in the wanted direction is the criterion.

206. As—

$$W_0 = 4\pi r^2 \cdot \frac{E_0^2}{240\pi}$$

and—

$$W = \frac{E_0^2 r^2}{240\pi} \int \int f^2(\theta, \varphi) \cos \theta \, d\theta \, d\varphi.$$

Then, $G = W_0/W$

$$G = \frac{4\pi}{\int \int f^2(\theta, \varphi) \cos \theta \, d\theta \, d\varphi} \text{ or } W = \frac{E_0^2 r^2}{60G}$$

so that the power gain G is determined in principle by the function $f(\theta, \varphi)$ which represents the polar diagram. Thus if $f(\theta, \varphi)$ is completely known from calculation, or by experiment, then G , in principle, can be obtained from the formula above.

207. Some typical values for the gain of simple aerials are:—

(1) Hertzian oscillator—($E = E_0 \cos \theta$)
 $G = 1.5$

(2) half-wave aerial—

$$\left(E = E_0 \frac{\cos\left(\frac{\pi}{2} \sin \theta\right)}{\cos \theta} \right). \quad G = 1.63$$

FACTORS GOVERNING MAXIMUM RANGE OF RADAR EQUIPMENTS

Noise

211. In any radar equipment with a display tube of the conventional type in which the echoes appear as deflections of the trace, the trace is subject to random “noise” fluctuations along its entire length.

212. These disturbances originate partly in the aerials and partly in the receiver, and their amplitude sets a limit to the range at which a target can be detected, since if an echo becomes too small it will be lost amongst them.

(3) Uniformly illuminated aperture (large enough to give marked beaming) of area A $G = 4\pi A/\lambda^2$

Maximum field strength at range r

208. As—

$$W = \frac{E_0^2 r^2}{60G}$$

$$E_0^2 = \frac{60WG}{r^2} \quad \text{where } rE_0 \text{ is in volts, and } W \text{ in watts.}$$

Then, knowing the power gain G of the aerial system, one can easily find the field strength E_0 produced at range r when power W is radiated from it.

209. Note, also that field E_0 , and therefore the field $E = E_0 f(\theta, \varphi)$ in a specified direction, is inversely proportional to the distance r from the source whatever the nature of $f(\theta, \varphi)$. As, for example, the maximum field strength E_0 at a distance of 10 kilometres from a half wave aerial radiating a power of 60 kilowatts is calculated as follows—

If $G = 1.63$; $r = 10^4$ metres; $W = 6 \cdot 10^4$ watts

$$E_0^2 = \frac{60 \times 60 \times 10^3}{10^4} \times 1.63$$

$$E_0 = 60\sqrt{0.163} = 24 \text{ volts per metre.}$$

Receiver and aerial noise

210. The value of the power gain G of an aerial in transmission can also be employed to compare the performances of aerials in reception. It may be deduced from the reciprocity theorem that the powers delivered by different aerial systems in reception and into matched loads, are proportional to their power gains G when the same plane electromagnetic wave is incident on them along the directions of best reception.

213. For this reason it is usual to measure the amplitude of the signal received from a target in terms of the noise amplitude, and to express it as a signal-to-noise ratio, S/N . Thus,

$$S/N = \frac{\text{Amplitude of the signal appearing on the tube}}{\text{Amplitude of the noise fluctuations on the tube}}$$

$$= \frac{\text{Signal voltage applied to the tube}}{\text{Noise voltage applied to the tube}}$$

$$= \sqrt{\frac{\text{Signal power}}{\text{Total noise power at input}}}$$

214. It is conventionally supposed that when the signal-to-noise ratio is unity the target is at maximum visible range, and that when the amplitude of the signal is less than that of the noise, the echo can no longer be detected.

215. Since the noise voltage sets a final limit to the range of the station it is advisable to enumerate various sources of noise, and to see the effect of reducing the total noise. These are as follows:—

216. *Noise due to dry joints.*—Dry joints and faulty connections both in the aerial and feeder systems and in the receiver will give rise to noise. This can obviously be remedied and need not be further discussed.

217. *Thermal agitation or Johnson noise.*—This is due to the thermal agitation of the electrons in the material of the conductors. When a current flows through a conductor this thermal agitation sets up between its ends a random voltage variation, the magnitude of which depends on its resistance and its absolute temperature. The noise voltage produced in this way in the HF stages of a receiver is independent of the frequency to which the receiver is tuned, but is proportional to the width of the frequency band to which the receiver is sensitive.

218. *Shot noise.*—Noise is produced in valves owing to the “shot effects”, or uneven emission of electrons from the filament. It occurs at all frequencies and cannot be wholly eliminated. Like the noise produced by thermal agitation in the conductors, it is proportional to the bandwidth over which the receiver is working.

219. *Partition noise.*—In a screen-grid valve some electrons are picked up by the screen and some by the anode, and the ultimate destination of any individual electron is purely a matter of chance. This introduces an extra random element in the arrival of electrons at the anode and gives rise to a new kind of noise. Unlike the two previous types of noise it not only depends on the bandwidth of the receiver, but also on the absolute frequency, and increases as the frequency becomes higher. Attempts have been made to reduce the partition noise by specially designed valves, of which two types may be briefly mentioned here.

- (1) In the beamed tetrode and pentode valves the effect is reduced by beaming the electrons through the screen grid.

It would also be possible to eliminate this effect by using a triode, but unfortunately in an ordinary triode the anode/grid capacity causes feedback which renders the use of such a valve impossible at high frequencies.

- (2) In the grounded grid type of triode, however, this difficulty is overcome by earthing the grid and applying the signal to the cathode. This type of valve is often used in the first stages of amplification.

220. *Aerial noise.*—Some noise will always occur in the aeriels. It is thermal in origin and is present at all frequencies. Its magnitude depends on the bandwidth of the receiver. A second type of aerial noise may also appear. It is due to atmospheric disturbances of various kinds, and is much less pronounced at higher frequencies.

221. The various sources of noise may be considered as natural “jamming” sources of power that tend to mask the legitimate signal. Since the fluctuations of noise potential are random, the noise powers may be added to give the total noise power. It is customary to pretend that all sources of noise operate at the input grid of the first valve, and the equivalent total noise power is usually estimated on this basis. The limit of detection may be assumed to have been reached when the signal power delivered by the aerial is equal to the total noise power. The signal-to-noise ratio is then unity.

Pulse and band width

222. In an infinite train of sinusoidal oscillations of constant frequency f cycles per second and of uniform amplitude, only one frequency f is present. If, however, the amplitude of the waves varies periodically the wave train can be analysed into two or more trains of different amplitudes and frequencies. In other words any amplitude modulated wave may be regarded either as a single wave train of fixed frequency and varying amplitude or as a series of superimposed wave trains, each of which has uniform amplitude and frequency, although its frequency differs from that of the other wave trains in the series.

223. In radar there is a particular application of this principle, since the transmitter is radiating not continuously but in short pulses. Suppose that a radar pulse consists of a burst of waves of frequency f . This pulse can no longer

be considered as a train of waves of the single frequency f as it could if its duration were infinite; the very fact that it continues for a finite time causes it to contain frequencies on either side of the frequency f , and it is possible to analyse it into an infinite number of superimposed wave trains whose frequencies vary from some lower limit ($f - \Delta f$) to an upper limit ($f + \Delta f$) as indicated in fig. 45. The frequency f is known as the *carrier frequency* and the total range of frequency, $2\Delta f$, is known as the *band width*.

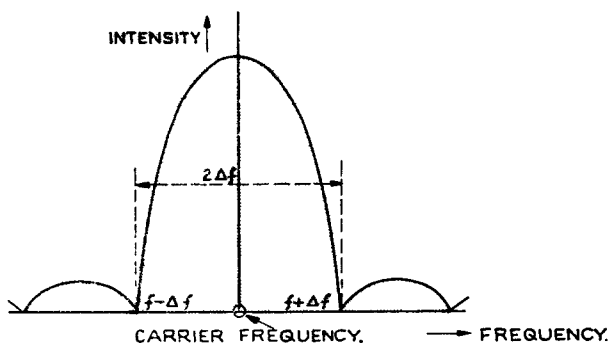


Fig. 45—Power/frequency graph

224. The band width of a pulse is determined by its duration, and the band width associated with any steep-sided pulse measured in megacycles is of the same order of magnitude as the reciprocal of the duration of the pulse in microseconds. Thus a one-microsecond pulse has a band width of the order of one megacycle, while a ten microsecond pulse has a band width of the order of 100 kilocycles.

225. The band width associated with a pulse also depends on the steepness of its sides, a steep-sided pulse having a wider band than one of the same duration whose sides are not so steep.

226. The earliest types of radar equipment were designed primarily to give early warning of the approach of enemy aircraft. In such equipments long range was of primary importance, and it was usual to employ a fairly wide pulse (10 to 15 microseconds, for example, on CH stations), so that the band width was small. Such stations fulfilled their purpose very well indeed, but it is impossible to use them for precision work for two reasons:—

- (1) The pulse was so wide that the systems had very low resolving power. The echoes from two aircraft at almost the same range would be seen superimposed on the tube.

- (2) The side of the pulse was not sufficiently steep to give great accuracy in range finding.

227. In order to overcome these two disadvantages, it is necessary in precision equipment to use pulses which are both narrower and steeper sided. This means increasing the band width of the pulse. To obtain an exact reproduction of the pulse shape on the display tube, and for the echoes to be well defined, the receiver must be sensitive to the whole range

of frequencies carried by the pulse. Thus increasing the accuracy and discrimination means increasing the band width of the receiver. This in turn means increasing the noise, since the amplitude of the noise depends on the bandwidth. Referring back to para. 214, it is seen that the increase in noise decreases the maximum range. Thus extra accuracy and resolving power can only be obtained at the expense of range.

228. It is clear then that in designing radar equipment the opposing claims of range on the one hand and accuracy and resolving power on the other hand must constantly be borne in mind. In early warning equipments such as CH and CHL, accuracy is sacrificed for long range, while in equipments such as GL and AI the maximum range is very short, but ranges can be measured with high precision.

Maximum range of detection

229. Consider the aerial system of a beamed radar device comprising a transmitting aerial whose power gain is G_T and a receiving aerial with power gain G_R . Let the aerials be directed so that an aircraft at range r miles lies on the axes of the aerial beams. When power W_T watts is transmitted the field strength E_A in the wave incident on the aircraft is given by—

$$E_A^2 = \frac{60W_T G_T}{r^2}.$$

This field excites oscillating currents in the aircraft which reradiates a scattered wave. The scattered wave has a polar diagram, consequently, the field strength in the scattered wave that reaches the receiving aerial depends on the aspect presented by the aircraft to the radar equipment.

230. Suppose that the aircraft is always approaching the station, so that a head-on aspect is presented. The field strength E_R at the receiver, of the scattered wave, is proportional to $\frac{E_A}{r}$. Then—

$$E_A = \frac{sE_R}{r}$$

where the coefficient of proportionality, s , is the aircraft scattering factor for the head-on aspect.

231. The flux of power in the scattered wave, at the receiver is $E_R^2/240\pi$ watts per sq. metre, and the power collected by the receiving aerial is proportional to this flux and to the equivalent absorbing area A_R of the aerial. If we put $G_R = \frac{4\pi A_R}{\lambda^2}$ then the power absorbed is proportional to $\lambda^2 G_R$.

232. Thus, the power W_R delivered to the receiver by its aerial system—

$$\begin{aligned} W_R &\propto \lambda^2 G_R E_R^2 \\ &\propto \frac{s^2 G_R E_A^2 \lambda^2}{r^2} \\ &\propto \frac{s^2 G_R G_T W_T \lambda^2}{r^4} \end{aligned}$$

233. Let the equivalent noise power introduced in effect at the receiver input be W_N . Then the limit of detection is assumed to be reached when $W_R = W_N$, the signal-to-noise ratio then being unity. Note that it is possible on occasions to "see" the signal inside the noise with signal-to-noise ratios somewhat less than unity, but it is convenient to adopt the above criterion for the operational limit to detection.

234. The maximum range of detection $r_{\max}$ then occurs when $W_R = W_N$ with the aircraft in the beams of both aerals. It follows that, where $r_{\max}$ is the maximum possible range of detection for an aircraft in head aspect with the scattering factor s ,

$$r_{\max} \propto \left(\frac{W_T}{W_N} \right)^{0.25} (G_T G_R)^{0.25} (s\lambda)^{0.5}$$

This expression enables us to discuss the influence of the respective parameters in maximum range of detection.

Transmitter power

235. If, keeping other factors unchanged, an attempt is made to increase range by increasing transmitter power, then $r_{\max}$ increases in proportion to the fourth root of the transmitter peak power. For instance, in order to double the range the transmitter power must be increased by a factor of sixteen. Consequently, it is seldom possible to achieve a spectacular improvement in range by this means.

Noise power of receivers

236. A reduction in noise power W_N has the same result as the same proportional increase in transmitting power. Consequently considerable improvements in over-all performance are obtained when attention is paid to the correct design of the receiver system to render the noise power W_N as small as possible.

237. Suppose that by correct design of input circuits and of first valve the noise is made as small as practicable; yet it still remains to determine the most suitable band width for the receiver. Since the noise power W_N is proportional to the band width, an improvement in signal-to-noise ratio can be made by decreasing the band width of the receiver. The pulse, however, possesses a spectrum whose width in cycles per second is given by $2/T$ where T is the duration of the pulse in seconds, so that if the band width of the receiver is made too narrow the pulse shape will be impaired by rejection of important frequency components.

238. To restore the pulse shape the pulse is lengthened so that its spectrum falls within the receiver band width. The precise balance struck between band width and pulse length is determined by the particular operational use for which the radar equipment is designed.

239. In early warning equipments such as CH it is important to make $r_{\max}$ as great as possible. Consequently, a large signal-to-noise ratio at long ranges is the design criterion. Such equipments, therefore, employ narrow band receivers and transmit relatively long pulse lengths, whose narrow spectra fall within the band width. Offsetting the gain in long-range detection, however, there is a loss in resolving power in range, since the long pulses

occupy lengths of 2 or 3 miles of the time base. Thus range accuracy and resolving power, which is the ability to distinguish echoes at nearly equal ranges, are poor. Thus long-range detection and good range discrimination are mutually exclusive in a single equipment.

240. In equipments designed to assist gun-laying, long-range detection is less important than accurate measurement of range. In such equipments steep narrow pulses are radiated and wide band receivers are employed. The noise power is, therefore, increased and $r_{\max}$ reduced.

241. It is technically simpler to produce short and steep-sided pulses on wavelengths of several centimetres than on the longer radar wavelengths, consequently centimetre wave equipments are superior to earlier equipments where accuracy of range as well as of angular measurement is important.

Scattering factor

242. The scattering factor depends on the wavelength of operation compared with the aircraft dimensions, and also, as already mentioned, on the aspect of the aircraft. In general, large aircraft such as four-engined bombers give stronger returns than small aircraft such as fighters. It is roughly true that when the wavelength is very small compared with the aircraft dimensions, as with centimetre waves, the scattering factor s is proportional to $\frac{1}{\lambda}$, and $s\lambda$ is therefore independent of the wavelength λ .

243. When the wavelength λ is about double the span of the aircraft then s is roughly proportional to λ , and when λ greatly exceeds the span s is roughly proportional to λ^{-4} .

Aerial power gain

244. We suppose that our aerials are represented by equivalent uniformly radiating and receiving areas A_T and A_R .

Then:—

$$G_R = \frac{4\pi A_T}{\lambda^2} \text{ and } G_R = \frac{4\pi A_R}{\lambda^2}$$

and

$$r_{\max} \propto \left(\frac{W_T}{W_N} \right)^{0.25} \cdot (A_T A_R)^{0.25} \cdot \left(\frac{s}{\lambda} \right)^{0.5}$$

Thus with aerials of fixed dimensions the maximum range at centimetre wavelengths increases as λ is reduced, since s does not change rapidly in this region.

245. At a fixed wavelength, using centimetre wavelengths, the maximum range increases with aerial surface. The formula shows that if the geometrical forms of the areas A_T and A_R are preserved while their linear dimensions are increased, then $r_{\max}$ is directly proportional to the linear dimensions of the aerials. Consequently, it is more profitable, when space is limited, to use a single large aerial with common T/R, than two separate aerials.

246. The use of common T/R is the universal practice adopted for airborne centimetre wave equipments. Since the single aerial in common T/R has at least double the diameter of the two separate aerials that could be installed in the same radome (perspex cover for a scanner), the range $r_{\max}$ of the equipment is double at least.

247. In centimetre wave ground equipments, such as GL, separate aerials are sometimes employed for transmission and reception.

CHAPTER 3

TRANSMISSION LINES AND WAVEGUIDES

LIST OF CONTENTS

	Para.		Para.
Introduction	1	Elimination of standing wave on line by stub	37
Steady electric field	3	Line with loss	38
Steady magnetic field	6	RF cables	39
Transmission lines		Loss in coaxial and transition to waveguides	40
Parallel strip transmission line	8	Approach to rectangular guide	41
Relation between E and H	9	Synthesis of H wave in rectangular guide ...	42
Characteristic impedance	11	Relation between wavelength in space and	
Power carried by wave	12	in the guide	44
Resistance of a thin film	14	Cut-off in a rectangular waveguide	47
Non-reflecting termination to a transmission		Circular waveguides	
line	15	H ₁₁ wave in circular pipe	50
Characteristic impedance of coaxial and		E ₀₁ wave in circular pipe	51
twin wire lines	16	Currents in waveguides	
Short-circuited transmission line	18	Currents in walls of rectangular guide	
Short-circuited line: special cases	21	carrying H ₀₁ wave	53
Termination of a transmission line by any		Slots and joints in rectangular guide with	
resistance	24	H ₀₁ wave	55
Input impedance to a line terminated in a		Series and shunt side arms	58
resistance	28	Waveguide technique	
Transmission line calculator	29	Iris in rectangular guide with H ₀₁ wave	59
Examples on the transmission line		Rotating joint using H ₀₁ to E ₀₁ transformer	64
calculator	32	Bends	66
Further examples on the transmission		Launching waves in guides	67
line calculator	33	Common T and R	
Deduction of terminating impedance from		Use of common T and R	69
standing wave measurements	34	1½ metre common T and R	70
Use of admittances	35	Waveguide common T and R	73

LIST OF ILLUSTRATIONS

	Fig.		Fig.
Coaxial and twin wire line	1	Line terminated in resistance	23
Waveguides	2	Voltage on resistance-terminated line ...	24
Electric field in condenser	3	Quarter-wave transformer	25
Electric field near metal plate	4	Circle of impedance	26
Magnetic field in solenoid	5	Impedance circles and "n" arcs	27
Single turn coil	6	Determination of input impedance	28
Magnetic field near metal plate	7	Shunt elements	29
Parallel strip line	8	Stub match for 3:1 standing wave	30
Fields in strip line	9	Coaxial feeder	31
Section of strip line	10	Fields with inner removed	32
Relationship between vectors for plane wave	11	Circular waveguide	33
Measurement of thin film resistance	12	Magnetic fields on strip lines	34
Termination of strip line	13	Rectangular guide	35
Effect of λ/4 short-circuit	14	Cross section of guide	36
Field in coaxial line	15	Synthesis of H ₀ wave	37
Element of resistive film	16	Fitting side walls	38
Fields in twin lines	17	Guide with H ₀₂ mode	39
Standing wave showing magnetic field	18	Relationship between λ _g and λ _a	40
Standing wave showing electric field	19	Rays of downgoing wave	41
Line less than λ/4	20	Rays of upgoing wave	42
Line between λ/4 and λ/2	21	Bouncing wave effect	43
Relationship between currents in opposite		Propagation well below cut-off	44
waves	22	Propagation at cut-off	45

	<i>Fig.</i>		<i>Fig.</i>
Distortion from rectangular to circular guide	46	Resonant iris	65
H_{11} wave in circular pipe	47	Captive screw	66
Azimuthal variation H_{11} mode	48	H_{01} to E_{01} transformer	67
E_{01} wave	49	Field in transformer	68
Currents in walls of guide	50	Rotating joint	69
Transmission line with stubs	51	Field in circular guide with E_{01} mode	70
Slots in walls of guide	52	Variation of output plane	71
Flanges	53	H_{11} field	72
Bolting of flanges	54	Ring filter	73
Choke joint	55	Right angle bend	74
Series arm	56	Launching of H_{01} wave	75
Analogue of series arm	57	Common T and R aerial	76
Shunt arm	58	$1\frac{1}{2}$ metre common T and R switching	77
Analogue of shunt arm	59	Rhumbatron	78
Capacitive iris	60	Loop coupling	79
Field near iris	61	Window coupling	80
Analogue of capacitive iris	62	T-R cell	81
Inductive iris	63	Waveguide T-R	82
Analogue of inductive iris	64	Equivalent circuit of waveguide T-R	83

CHAPTER 3

TRANSMISSION LINES AND WAVEGUIDES

INTRODUCTION

1. At radio frequencies, when the wavelength is fairly small, wires connecting different parts of a circuit must be short. If this is not so, the connecting wires themselves become circuit elements and may also give trouble by radiating. When it is not possible to make connecting wires short—as, for example, in the connection between transmitter and aerial or between aerial and receiver—a specially designed connection has to be used which will not radiate appreciably and which will join up the two

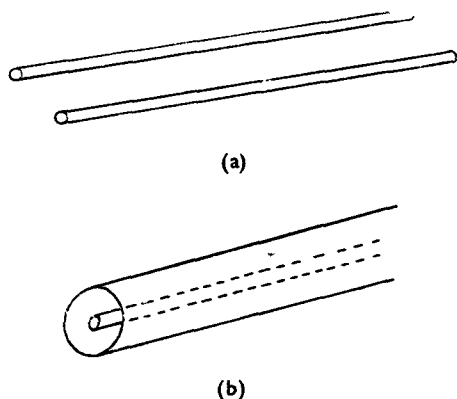


Fig. 1.—Coaxial and twin wire line

parts of the system as if they were effectively close together. Such long wire connections are called transmission lines and take the form either of a twin wire line or a coaxial line (see fig. 1). At centimetre wavelengths, a waveguide may be used; this is a hollow pipe, usually of rectangular cross-section but occasionally circular, see fig. 2.

2. In considering the transmission of power by a transmission line, it is simplest to regard the line as having distributed inductance and capacity and to treat the problem from the circuit or filter view point, working out the voltages and currents at points along the line. Waveguides cannot be easily treated in this way as there are no obvious “go” and “return”

paths in a wave guide, and the waveguide has to be investigated from the point of view of an electromagnetic field which travels up the inside of the pipe. However, transmission lines can also be regarded as a kind of “waveguide” forming a track or rails along which power is guided from one point to another in the form of an electromagnetic wave. In order to achieve a uniform treatment of the subject, both lines and guides will be investigated as far as possible from the field point of view. Since, however, circuit concepts are more familiar than field concepts, they are used as required to help the argument when the field method is difficult.

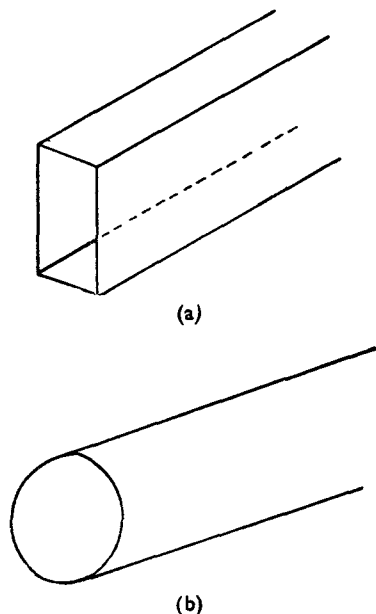


Fig. 2.—Waveguides

Steady electric field

3. The steady electric field is most familiar in the case of a parallel plate condenser whose metal plates are connected to a battery. This is shown in fig. 3. The plates are taken to be

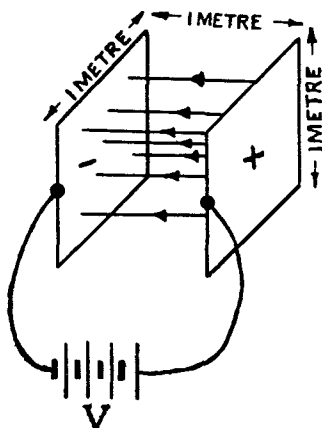


Fig. 3.—Electric field in condenser

one metre square and one metre apart. The battery voltage is V volts. Edge effects are neglected; if necessary, guard rings may be put round the edges to eliminate distortion of the field. The difference of potential between the plates is V volts and there is a charge q coulombs on each plate given by

$$q = K_0 V$$

where K_0 is the capacity of the condenser in farads. Proceed to calculate K_0 thus:—

$$\begin{aligned} K_0 &= \frac{\text{area of either plate}}{4\pi \times \text{spacing}} \\ &= \frac{10^2 \times 10^2}{4\pi \times 10^2} \\ &= \frac{10^2}{4\pi} \text{ centimetres (or electrostatic units)} \end{aligned}$$

$$1 \text{ microfarad} = 900,000 \text{ centimetres}$$

$$\text{Hence } 1 \text{ farad} = 9 \times 10^{11} \text{ centimetres.}$$

$$\therefore K_0 = \frac{10^2}{4\pi \times 9 \times 10^{11}} = \frac{10^{-9}}{36\pi} \text{ farads.}$$

This enables one to obtain the charge on the plates in terms of the voltage.

4. Now consider the space between the plates to be filled with electric field. Electric lines of force start on the positive plate and go over to the negative plate. Since the difference of potential between the plates is V volts and their distance apart is one metre, the electric field, E , is given by

$$E = V \text{ volts/metre.}$$

We can now relate the electric field between the plates with the charge on the plate

$$q = K_0 E$$

$$E = q/K_0 \text{ volts/metre.}$$

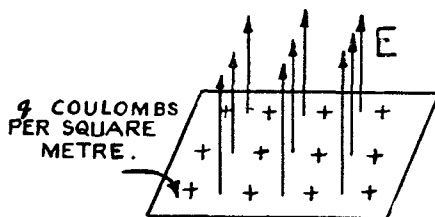


Fig. 4.—Electric field near metal plates

Finally one notes that the electric field is perpendicular to the metal plates.

5. Now consider a large flat metal sheet, not necessarily part of a parallel plate condenser. Let there be charge on the surface of the metal sheet. Then we take over the results of the condenser and say that there must be an electric field E standing perpendicular to the sheet (see fig. 4). If the surface density of charge is q coulombs/square metre, then the electric field is given by

$$E = q/K_0 \text{ volts/metre}$$

where K_0 has the value $10^{-9}/36\pi$. After the lines leave the surface of the metal sheet they may bend over or behave in any manner, but in the immediate neighbourhood of the sheet they must stand perpendicular to it.

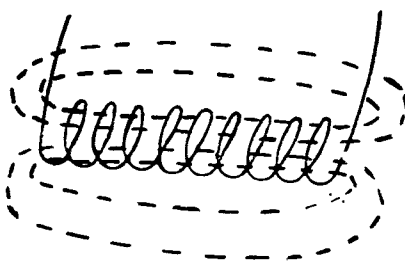


Fig. 5.—Magnetic field in solenoid

Steady magnetic field

6. The steady magnetic field is most familiar in the case of a solenoid with current passing through its turns. If the solenoid is fairly long, the magnetic field is appreciable only inside the solenoid. Outside, the lines are sparsely distributed in space, while inside there is a uniform steady magnetic field (see fig. 5). The strength of this magnetic field, H , depends on the amperes flowing and on the number of turns per unit length, i.e.

$$H = In/l$$

where I is the current, n the total number of turns and l the length of the solenoid. In order to get a more convenient expression for the magnetic field, consider a solenoid one metre

long, consisting of a single "turn" and carrying a current i amperes (see fig. 6). Then the magnetic field, H , inside the solenoid is given by

$$H = i \text{ amperes/metre.}$$

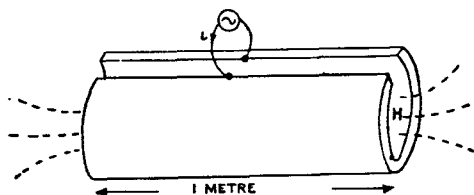


Fig. 6.—Single turn coil

7. Note also that the magnetic field lies parallel to the surface of the metal in the inside. If we now imagine the metal rolled out flat we obtain a strip one metre wide carrying a surface current i amperes. Above the sheet is a magnetic field H equal to i amperes/metre (see fig. 7). Generally, if we have a large metal sheet carrying a current on its surface, then there will be a magnetic field lying parallel to the sheet and perpendicular to the direction of the current. If the current per metre strip of sheet is i amperes/metre then the magnetic field is i amperes/metre.

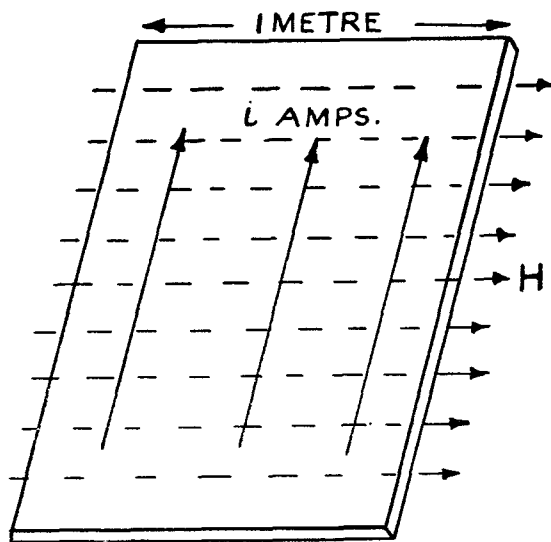


Fig. 7.—Magnetic field near metal plate

TRANSMISSION LINES

Parallel strip transmission line

8. The fields in the twin wire and coaxial transmission lines are a little complicated and it is convenient to consider initially a more simple type of transmission line in the form of two parallel metal strips (see fig. 8). The height of a strip is b metres and the distance apart is a metres. An A.C. generator is attached



Fig. 8.—Parallel strip line

at the beginning of the strips and pushes, alternately, positive and negative charges on to the strips. These charges appear to move down the strips with the speed of light (i.e. 3×10^8 metres/second). It is, of course, unlikely that charges could move as fast as that. Rather it would seem that the electrons in the thickness of the strips are sucked out to the surface or repelled as required so as to give the appearance of charges moving along the strips. If the line stretches a long distance from the generator so that one is not worried about the ultimate destination of the charges, the position, after a little time, will be as shown in fig. 9 where one is supposed now to be looking down on the edges of the strips.

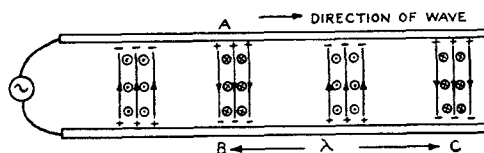


Fig. 9.—Fields in strip line

One can then fill in the associated electric and magnetic fields. The electric field arises from the charges on the surfaces of the strips and is, of course, perpendicular to the surface. The magnetic field arises from the currents on the surfaces of the strips. "Current" really means "rate of movement of charge". All the charges, are moving to the right. A positive charge moving to the right means an electric current directed to the right. A negative charge

moving to the right has to be considered as equivalent to a conventional current moving to the left. One thus obtains magnetic fields perpendicular to the plane of the paper and directed as shown in fig. 9. The electric field is in the plane of the paper. The whole pattern moves to the right with the velocity of light. The distance BC between points in the same phase is a wavelength.

Relation between E and H

9. Since the electric field, E , arises from the charges and the magnetic field, H , from the motion of the charges, one sees that E and H are related. If E is increased, H increases in proportion. The fact that E and H cannot have unrelated arbitrary values is an important property of an electromagnetic wave. In order to find the relationship, take a short length of transmission, such as the portion AB in fig. 9,

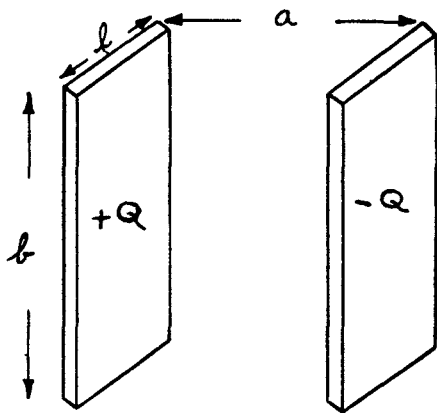


Fig. 10.—Section of strip line

of length l metres. This is shown in fig. 10. Let the charge, at any instant on one of the plates be Q coulombs. The charge, q , per square metre is

$$q = Q/(bl) \text{ coulombs/metre}^2.$$

Hence, the electric field E is given by

$$\begin{aligned} E &= q/K_0 \\ &= Q/(bK_0l) \\ &= \frac{Q}{bl} \times \frac{36\pi}{10^9} \text{ volts/metre.} \end{aligned}$$

The current, I , along the strip is the charge passing in one second. The speed of movement of the charge is 3×10^8 metres/second. Time for charge to move length l metres is

$$l/(3 \times 10^8) \text{ seconds.}$$

$$\text{Hence, } I = \frac{Q \times 3 \times 10^8}{l} \text{ amperes.}$$

This current is spread over a strip of width b metres. The current, i , per metre strip is

$$i = \frac{Q \times 3 \times 10^8}{b \times l} \text{ amperes/metre.}$$

Hence, the magnetic field H , is given by

$$H = \frac{Q}{bl} \times 3 \times 10^8 \text{ amperes/metre.}$$

Dividing, one finds

$$\begin{aligned} E/H &= 120\pi \\ &= 377 \text{ (approx.).} \end{aligned}$$

The ratio E/H has the dimensions volts/amperes, i.e. ohms. So, the number 120π is dimensionally ohms. The equation is analogous to the circuit relation

$$V/I = Z$$

and therefore 120π ohms is sometimes called the "wave impedance".

Denoting it by Z_w , we can write

$$E/H = Z_w$$

with $Z_w = 120\pi$ ohms.

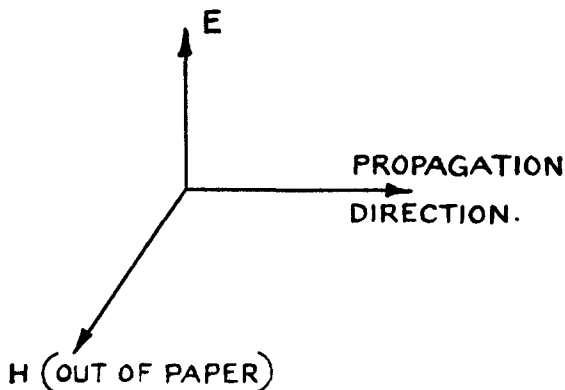


Fig. 11.—Relationship between vectors for plane wave

10. This relation is also true for any simple plane wave, such as the wave at a large distance from a transmitting aerial. Note that E and H are at right angles to one another and to the direction of propagation (see fig. 11). A right-handed corkscrew rotation of from E to H would make it advance in the direction of propagation. Also E and H oscillate in phase, reaching their maximum values at the same time and place.

Characteristic impedance

11. The relation between E and H , namely, $E/H = 120\pi$ does not involve the dimensions of the strip transmission line. If, however, we consider V and I the dimensions come in:—

$$\begin{aligned} V &= E \times a \text{ volts} \\ I &= H \times b \text{ amperes} \\ \frac{V}{I} &= \frac{E}{H} \times \frac{a}{b} \\ &= 120\pi \times \frac{a}{b} \text{ ohms.} \end{aligned}$$

This is called the characteristic impedance of the line and is usually denoted by Z_0 . It is the ratio of volts across the line to current flowing at the point, when a single wave is travelling up the line.

Power carried by wave

12. Since the power is not supposed to leak away, it will be sufficient to calculate it at the input to the line. We then have for a strip line, using r.m.s. values,

$$\begin{aligned} \text{Power} &= V \times I \\ &= E \times a \times H \times b \\ &= EH \times ab \text{ watts.} \end{aligned}$$

13. We regard the power as carried by the wave and since ab is the cross-sectional area of the wavefront, we see that the power, S , carried per unit area of wavefront is

$$S = EH \text{ watts/square metre.}$$

This can also be expressed in the form

$$S = E^2/Z_w \text{ watts/square metre.}$$

Sometimes S is called the *Poynting vector*, after Poynting who discovered it. The energy is carried by the magnetic field and by the electric field. It can be shown that the electric and magnetic energy in the plane wave are equal.

Example:—A transmission line consists of two parallel metal strips 3 cm. wide and 1 cm. apart. If the spark over field is 30,000 volts/cm. what is the maximum power the line will carry?

$$\begin{aligned} E_{\text{peak}} &= 30,000 \text{ volts/cm.} \\ &= 3 \times 10^6 \text{ volts/metre.} \\ E_{\text{rms}} &= 2.12 \times 10^6 \text{ volts/metre.} \end{aligned}$$

$$\text{Power per unit area} = E^2/Z_w$$

$$= \frac{2.12^2 \times 10^{12}}{120\pi} \text{ watts/metre}^2$$

$$\begin{aligned} \text{Area of cross-section} &= \frac{3}{100} \times \frac{1}{100} \text{ metre}^2 \\ &= 3 \times 10^{-4} \text{ metre}^2 \end{aligned}$$

$$\begin{aligned} \therefore \text{Power} &= \frac{2.12^2 \times 3 \times 10^{12} \times 10^{-4}}{120\pi} \\ &= \frac{4.5 \times 3}{1.2 \times \pi} \times 10^6 \\ &= 3\frac{1}{2} \text{ megawatts.} \end{aligned}$$

Resistance of a thin film

14. To terminate a transmission line with a resistance it is preferable to use a thin resistive film rather than an ordinary commercial resistor as used in circuit work. As we have seen, the wave travelling up the line extends over a considerable cross-section and, in order that it may intercept the whole of the wave, the resistance must extend over the same area. Such a resistance might be made up from

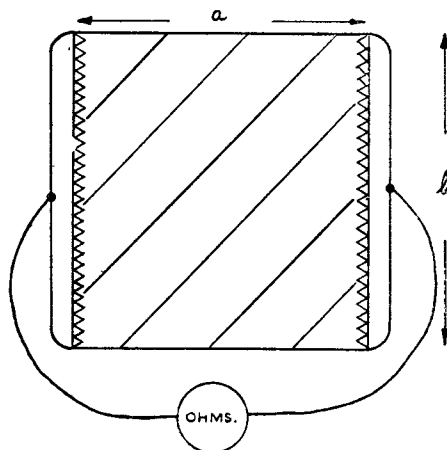


Fig. 12.—Measurement of thin film resistance

graphite or mica or cloth impregnated with aquadag. Take a thin resistive film of width a and height b . Attach long metal clips to the edges of the long sides as shown (see fig. 12) and connect to an ohm-measuring meter. Then the resistance R of the film will be directly proportional to a and inversely proportional to b . Thus,

$$R = \Omega \frac{a}{b} \text{ ohms}$$

where Ω is a constant. To find Ω , put $a = b$, i.e. choose a square film. Then

$$R = \Omega \text{ ohms}$$

so that Ω is the resistance per square of material.

Non-reflecting termination to a transmission line

15. On circuit arguments, the resistive termination which will absorb all the power travelling up a line is simple; it is the characteristic impedance, Z_0 ohms, of the line. The proof is well known and simple. On field ideas we must use a resistive film and an elaborate arrangement is required. A resistive film must be stretched across the line so as to intercept the whole of the wave. Also the line must be continued for a quarter of a wavelength and short circuited with a flat metal plate (see fig. 13). Currents flow in the film and

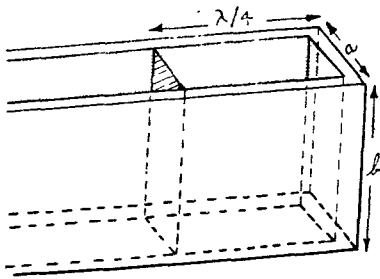


Fig. 13.—Termination of strip line

some wave leaks through into the quarter-wave section. Owing to the metal shorting-plate, the wave which leaks through goes up to the plate, is reflected and arrives back at the film. By the time it arrives back it has travelled $\lambda/2$ and the magnetic field of the reflected wave is in antiphase with the magnetic field of the primary wave at a point just behind the film*. There is, therefore, no magnetic field just behind the film, i.e. at D in fig. 14, which shows the strip transmission line view edge-on. At C we have the magnetic field H of the wave, coming up from the generator. Thus a current I flows in the strip of amount

$$I = Hb \text{ amperes.}$$

Similarly there is an electric field, E , across AB giving rise to a potential difference V of amount

$$V = Ea \text{ volts.}$$

If the film is to sustain this field system its resistance R must be such that

$$\begin{aligned} R &= V/I \\ &= \frac{E}{H} \times \frac{a}{b} \\ &= 120\pi \times \frac{a}{b} \text{ ohms.} \end{aligned}$$

* The action of the $\lambda/4$ short-circuit will be clearer later on.

Comparing, we see that its resistance Ω per square must be given by

$$\begin{aligned} \Omega &= 120\pi \text{ ohms per square} \\ &= 377 \text{ ohms per square.} \end{aligned}$$

Such a film will completely absorb the wave. The value of R is the same as Z_0 , as would be expected from the earlier arguments.

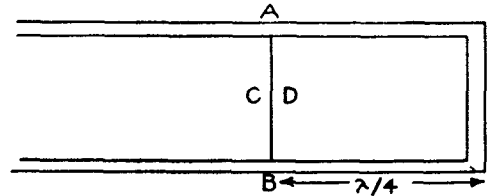


Fig. 14.—Effect of $\lambda/4$ short-circuit

Characteristic impedance of coaxial and twin wire lines

16. The field at any instant over a cross-section of a coaxial line is shown in fig. 15. The electrical lines, as usual, are perpendicular to the metal surfaces. The magnetic lines are parallel to the surfaces. As distinct from the parallel strip line, the fields are not uniform over the cross-section. The magnetic field is strongest near the inner conductor since the current density is greater on the inner than on the outer conductor. Similarly the electric field is stronger at the inner since the lines get further apart as they diverge radially. However, at any little part of the cross-section, such as indicated by the square in fig. 15, E and H are at right angles and the field in the little

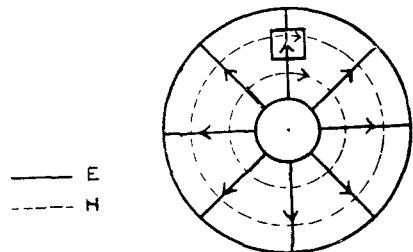


Fig. 15.—Field in coaxial lin.

square is similar to the field between two metal strips. Thus each element would be terminated without reflection by a 377-ohm film. The characteristic impedance is therefore obtained by placing a 377-ohm film over the end of the coaxial. A simple integration enables the value of Z_0 to be obtained. The resistance across the element bounded by circles of radii r and $r + dr$ (see fig. 16) is:—

$$120\pi \frac{dr}{2\pi r} \text{ ohms.}$$

Thus the total resistance is

$$Z_0 = 60 \int_{r_1}^{r_2} \frac{dr}{r}$$

$$= 60 \log_e \frac{r_2}{r_1}$$

where r_1 and r_2 are the inner and outer radii respectively. Since

$$\log_e N = 2.3 \times \log_{10} N$$

we find $Z_0 = 138 \log_{10} (r_2/r_1)$ ohms.

This applies for air dielectric.

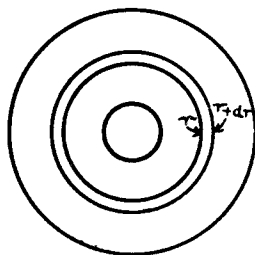


Fig. 16.—Element of resistive film

17. The case of the twin wire may be considered in the same way. The lines of electric and magnetic field at a cross-section are as shown in fig. 17. Here again the field is not uniform over the cross-section but, over any little area, E and H are at right angles and in the ratio 120π . A resistive film to absorb all the power would in theory have to extend to infinity since now the field is not confined. In practice the film might extend as shown in fig. 17. The calculation of Z_0 is in this case more difficult and only the result is quoted:—

$$Z_0 = 120 \log_e (2D/d)$$

where D is the spacing between the centres of the wires and d is the diameter of either wire. This can be rewritten

$$Z_0 = 276 \log_{10} (2D/d) \text{ ohms.}$$

In order to achieve a good absorption of the power the resistive films have to be used with a metal short-circuiting plate $\pi/4$ beyond the film.

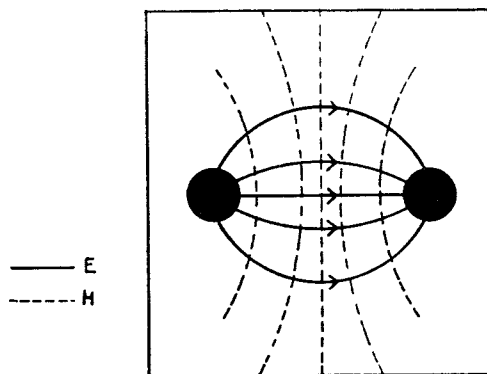


Fig. 17.—Fields in twin lines

Short-circuited transmission line

18. When a large metal plate sufficient to intercept all, or nearly all, of the wave is placed at the end of a transmission line, one has to consider the effect on the wave. The conditions at the surface of a good conductor are:—

E perpendicular to surface, i.e. $E_{\text{tangential}} = 0$
and H parallel to surface, i.e. $H_{\text{normal}} = 0$.

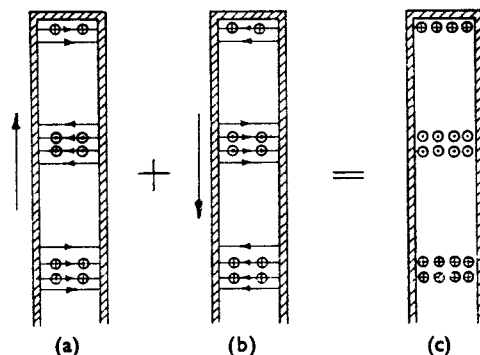


Fig. 18.—Standing wave showing magnetic field

19. Examine now the state of affairs at the short-circuit at the top of fig. 18(a). The electric field is lying tangential to the metal shorting plate: this is a state of affairs which cannot exist. The difficulty is overcome by reflection taking place—or, one may say, another wave is generated—and an equal wave travels down the line away from the short circuit as shown in fig. 18(b). This reflected wave is of such phase that its electric field is opposite to the electric field of the incident wave at the short-circuit. Now it is noticed

that when one wishes to reverse the direction of propagation of a wave one alters one of the quantities E or H but not both. If both E and H are reversed the direction of propagation is unaltered; it is, expressed simply, as if one moved to a point on the same wave distant $\lambda/2$ away. Hence, if E in the reflected wave is the reverse of E in the incident wave then H is the same in both waves. The situation is therefore as shown in fig. 18(b). When the two fields are added the result is as shown in fig. 18(c). Note that the field is wholly magnetic.

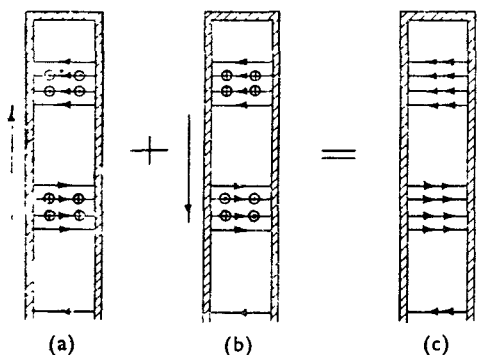


Fig. 19.—Standing wave showing electric field

20. One now considers the fields a quarter of a period later. In this time the incident wave has moved up $\lambda/4$ and the reflected wave down $\lambda/4$. The situation is therefore as shown in fig. 19, and in fig. 19 (c) it is seen that the field is now purely electric. The electric lines of force have appeared in between the positions of the magnetic lines in fig. 18(a). Half a period after the time of fig. 18, the field will be wholly magnetic but of opposite phase to fig. 18. Three-quarters of a period after the initial instant, the field will again be electric and so on.



Fig. 20.—Line less than $\lambda/4$

Short-circuited line: special cases

21. If the line is less than $\lambda/4$, then the fields obtained from figs. 18(c) and 19(c) at times $t = 0$ and $t = T/4$ are as shown in fig. 20. There is more magnetic energy than electric energy and the line tends to behave as though it were a coil; it is inductive.

22. If the line is between $\lambda/4$ and $\lambda/2$, the fields obtained from figs. 18(c) and 19(c) at $t = 0$ and $t = T/4$ are as shown in fig. 21. There is now more electric than magnetic energy and the line behaves like a condenser.

23. Finally if the line is $\lambda/4$ long exactly it will be seen that the electric and magnetic energies are exactly equal, the electric and magnetic fields are oscillating 90° out of phase and the line is equivalent to a parallel resonant L-C circuit.

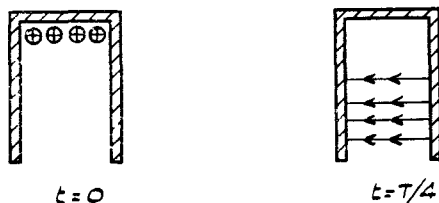


Fig. 21.—Line between $\lambda/4$ and $\lambda/2$

Termination of a transmission line by any resistance

24. When a transmission line is terminated in a resistance not equal to Z_0 , the fields or the voltage and current in a single travelling wave cannot be sustained by the resistance. As in the case of a short circuited line, a reflected wave is set up travelling in the opposite direction but not necessarily of the same amplitude as the incident wave. Since resistance is essentially a circuit concept we use V and I in this section rather than E and H . It will be recalled that in a travelling wave we have the relation

$$V/I = Z_0 \text{ ohms.}$$

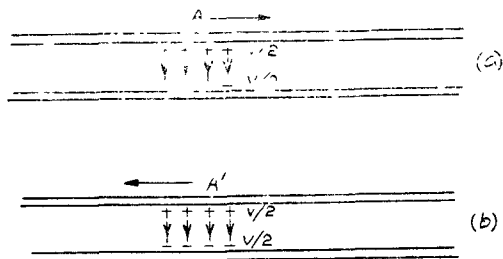


Fig. 22.—Relationship between currents in opposite waves

However, if the wave is travelling in the opposite direction, as in fig. 22(b), we must watch the sign of I . In fig. 22(a) at the point A the upper wire is at $+V/2$ volts and the current is to the right. In fig. 22(b) at A' , the upper wire is at $+V/2$ volts, but the current is now to the left

since the wave is in the opposite direction. If, therefore, we wish to combine currents when two such waves are superimposed, currents to the right may be regarded as positive and those to the left as negative. Then, the relation for the voltage V and current I in the reflected wave should be related by

$$V = -Z_0 I.$$

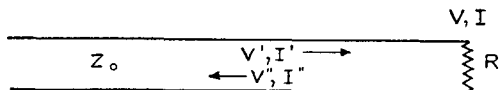


Fig. 23.—Line terminated in resistance

25. Consider now a line terminated in a resistance R ohms. As explained above, two waves will be required in order that the voltage and current at R will satisfy Ohm's law. Thus if V and I are the voltage and current at R (fig. 23)

$$V = V' + V''$$

$$I = I' + I''$$

and $V' = Z_0 I'$, $V'' = -Z_0 I''$.

Since Ohm's law must be satisfied

$$V = IR.$$

Hence, $R = \frac{V}{I}$

$$= \frac{V' + V''}{I' + I''}$$

$$= Z_0 \frac{(V' + V'')}{(V' - V'')}$$

Thus, $RV' - RV'' = Z_0 V' + Z_0 V''$

$$V''(R + Z_0) = V'(R - Z_0)$$

$$\frac{V''}{V'} = \frac{R - Z_0}{R + Z_0}$$

This formula gives the voltage reflection coefficient for a line of characteristic impedance Z_0 terminated in a resistance R .

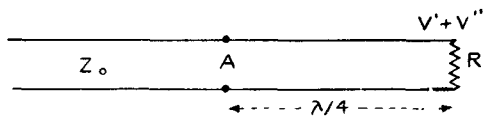


Fig. 24.—Voltage on resistance-terminated line

26. When $R = Z_0$ there is no reflected wave; the wave is completely absorbed in the resistance. When $R = 0$, i.e. a short circuit, then $V''/V' = -1$, indicating that the wave

is reflected with a reversal of phase of voltage. When R lies between 0 and Z_0 the wave is partly absorbed and partly reflected without reversal of phase of voltage. Similarly when R lies between Z_0 and ∞ , the wave is partly absorbed and partly reflected without change of phase of voltage. Consider then the latter case, $R > Z_0$. The voltage at R is the sum of V' and V'' . This is the maximum value the voltage on the line can have. Now measure back to A , a distance $\lambda/4$ from R (see fig. 24). The phase of the voltage of the incident wave at A is 90° earlier than the incident wave voltage at R . Also the phase of the voltage of the reflected wave at A is 90° later than the reflected wave voltage at R . Since V' and V'' are in phase at R they must be 180° out of phase at A , and the voltage at A is $V' - V''$. This is the lowest voltage on the line. The ratio of maximum voltage to minimum voltage is called the standing wave ratio and is denoted by s . Hence,

$$s = \frac{V' + V''}{V' - V''}$$

Using a previous relation this becomes

$$s = \frac{R}{Z_0}$$

27. This relation is very important. It says that if we measure the ratio of maximum to minimum voltage along the transmission line, and if there is a voltage maximum at the end of the line, the line is terminated in a resistance sZ_0 ohms. In the same way it is easy to argue that if there is a voltage minimum at the end of the line the termination is a resistance Z_0/s ohms.

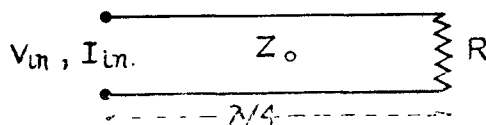


Fig. 25.—Quarter-wave transformer

Input impedance to a line terminated in a resistance

28. Let the line be $\lambda/4$ long and terminated in a resistance R (fig. 25). The voltage and current at R will be V_R , I_R given by

$$V_R = V' + V''$$

$$I_R = I' + I''$$

and $R = \frac{V_R}{I_R}$

At the input the voltages V' and V'' as well as the currents I' , I'' will now be in antiphase if in phase R , and in phase if in antiphase at R ,

$$\begin{aligned}\text{Thus } V_{in} &= V' - V'' \\ I_{in} &= I' - I''.\end{aligned}$$

$$\begin{aligned}\text{Then } Z_{in} &= \frac{V_{in}}{I_{in}} = \frac{V' - V''}{I' - I''} \\ &= Z_0^2 \frac{I' + I''}{V' + V''} \\ &= Z_0^2 \frac{I_R}{V_R} = \frac{Z_0^2}{R}.\end{aligned}$$

This is the "quarter-wave transformer" effect. The resistance is, so to speak, turned into its reciprocal or inverted with respect to Z_0^2 . Similarly it can be seen that if the line is $\lambda/2$ long, the voltages and currents are the same at the input as at the termination and then

$$Z_{in} = R$$

i.e. the half-wave line acts as a 1:1 transformer.

Transmission line calculator

29. When the line is terminated in a resistance but is not of length $\lambda/4$ or $\lambda/2$, the calculation of the input impedance is very difficult. Generally speaking the input impedance is complex, there is a resistive and a reactive term. If we have a small resistive termination R which is less than Z_0 , and if the line is less than $\lambda/4$ then the input impedance contains a resistive part and an inductive part. One can see it is inductive since in the limiting case of a short circuit the line would be inductive. As the length of the line is increased until it reaches $\lambda/4$, the impedance grows until it becomes Z_0^2/R (quarter-wave transformer) and is again non-reactive. Between $\lambda/4$ and $\lambda/2$ the impedance has capacitive reactance. This variation is shown in the transmission line calculator. (Refer to A.P.1093-E, fig. 185.)

30. The resistive axis is horizontal, positive reactances are measured up and negative reactances down on the vertical axis. In order that the calculator shall apply to any line, the resistances and reactances are supposed to be normalised, i.e. divided by Z_0 . Thus, if the terminating resistance is 25 ohms and the characteristic impedance 50 ohms we call the normalised termination

$$r = 25/50 = 1/2.$$

31. If the line is $\lambda/4$ long, then the normalised input impedance is $1/r$, i.e. 2; or 100 ohms, when multiplied by Z_0 again. Between lengths 0 and $\lambda/4$ the normalised input impedance lies

between 1/2 and 2 with a positive reactance. The locus of these input impedances lies on a circle as shown in fig. 26. Between $\lambda/4$ and $\lambda/2$ the reactance X is negative. Each dot on the circle indicates the length of line corresponding to the input impedance represented by the dot. In order that the calculator shall apply to any resistive termination, there are many circles. When dots corresponding to the same length of line are joined up they form arcs as shown in fig. 27. They are often called 'n' arcs, since they express the length of line in the form $n\lambda$.

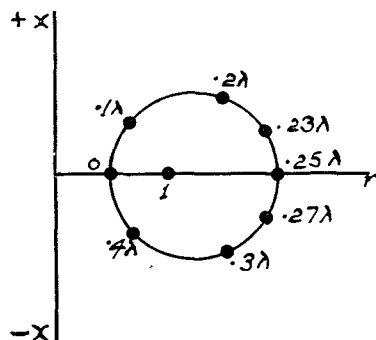


Fig. 26.—Circle of impedance

32. Examples on the transmission line calculator

Example.—A transmission line of characteristic impedance 600 ohms is terminated in a resistance of 300 ohms. The length of the line is 0.2λ . What is the input impedance?

First normalise the resistance:—

$$r = R/Z_0 = 300/600 = 0.5.$$

Find the point (0.5, 0) on the diagram, i.e. the point A in fig. 28. Travel round the circle

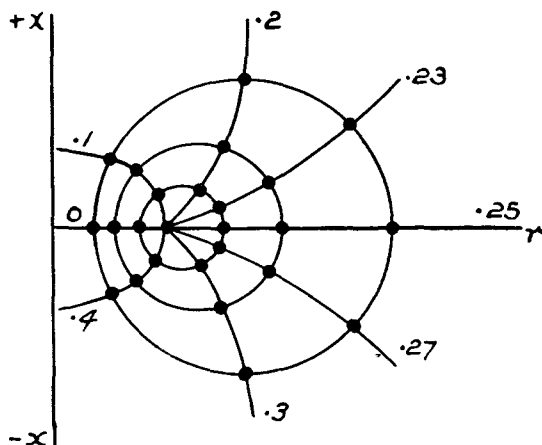


Fig. 27.—Impedance circles and "n" arcs

until the arc $n = .20$ is met, i.e. the point B and read off the answer:—

$$z_{in.} = 1.52 + j0.65.$$

This is normalised. The value in ohms is found by multiplying by Z_0 , i.e. by 600, and one finds

$$Z_{in.} = 912 + j390 \text{ ohms.}$$

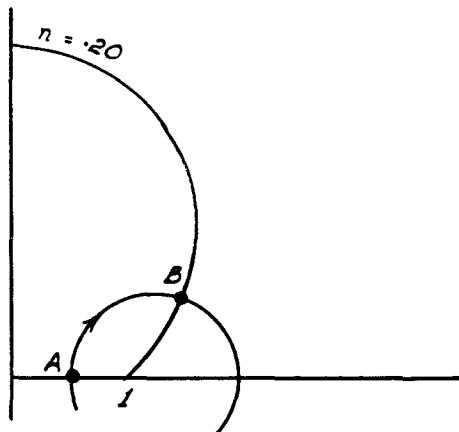


Fig. 28.—Determination of input impedance

Example.—A transmission line of characteristic impedance 600 ohms is terminated in a resistance of 300 ohms. Find the input impedance when the length of line is (a) $\lambda/8$, (b) $\lambda/4$, (c) $3\lambda/8$ and (d) $\lambda/2$.

$$r = R/Z_0 = 300/600 = 0.5.$$

Find the point (0.5, 0) on the diagram and travel round the circle until the arcs $n = 0.125$, 0.25, 0.375 and 0.5 are met, reading off the answers:—

$$(a) z_{in.} = .8 + j.6; Z_{in.} = 480 + j360 \text{ ohms.}$$

$$(b) z_{in.} = 2 + j0; Z_{in.} = 1200 + j0 \text{ ohms.}$$

$$(c) z_{in.} = .8 - j.6; Z_{in.} = 480 - j360 \text{ ohms}$$

$$\text{and (d) } z_{in.} = .5 + j0; Z_{in.} = 300 \text{ ohms.}$$

Note that $n = 0.5$ is the same as $n = 0$.

Example.—A transmission line of characteristic impedance 80 ohms is terminated in a resistance of 240 ohms. Find the input impedance when the length of the line is (a) $\lambda/8$, (b) $\lambda/4$ and (c) $3\lambda/8$.

$$r = 240/80 = 3.$$

Find the point (3, 0). Its n value is $n' = 0.25$. Add in succession .125, .25, .375 to n' and obtain the arcs $n'' = 0.375$, 0.5, 0.625.

Subtract 0, 5 as required and obtain $n'' = 0.375$, 0, 0.125. Go round the circle through (3, 0) until these arcs are met.

$$(a) z_{in.} = .6 - j.8; Z_{in.} = 48 - j64 \text{ ohms.}$$

$$(b) z_{in.} = .33 + j0; Z_{in.} = 27 \text{ ohms.}$$

$$(c) z_{in.} = .6 + j.8; Z_{in.} = 48 + j64 \text{ ohms.}$$

Example.—A transmission line of characteristic impedance 600 ohms is terminated in a resistance of 600 ohms in series with a condenser of capacity 9 pfs. The frequency is 30 Mc/s. What is the input impedance when the length of line is (a) $\lambda/8$, (b) $\lambda/4$ and (c) $3\lambda/8$?

$$X = -\frac{1}{2\pi f C} = -\frac{1}{2\pi \times 3 \times 10^7 \times 9 \times 10^{-12}} \\ = -600 \text{ ohms}$$

$$r + jx = \frac{R + jX}{Z_0} = \frac{600 - j600}{600} = 1 - j1.$$

Find the point (1, -1). Its n value is $n' = 0.34$. To this add 0.125, 0.25, 0.375 obtaining $n'' = 0.465$, 0.59, 0.715. Subtract 0.5 as required and find $n'' = 0.465$, 0.09, 0.215. Go round the circle through (1, -1) until these arcs are met.

$$(a) z_{in.} = 0.4 - j0.18; Z_{in.} = 240 - j108 \text{ ohms}$$

$$(b) z_{in.} = 0.5 + j0.5; Z_{in.} = 300 + j300 \text{ ohms}$$

$$(c) z_{in.} = 2.05 + j1; Z_{in.} = 1230 + j600 \text{ ohms}$$

33. Further examples on the transmission line calculator

Example.—A transmission line is $\lambda/8$ long, of characteristic impedance 300 ohms and short-circuited at its far end. What is its input impedance?

$$r + jx = 0 + j0$$

Find the point (0, 0). Travel round the "infinite circle" until the arc $n = 0.125$ is met.

$$z_{in.} = 0 + j1$$

$$Z_{in.} = j300 \text{ ohms (inductive stub).}$$

Example.—A transmission line is $\lambda/8$ long, of characteristic impedance 300 ohms and open-circuited at its far end. What is its input impedance?

$r + jx = \infty + j\infty$. "Find" the point (∞ , ∞). Its n value is $n' = 0.25$. Add 0.125 and obtain $n'' = 0.375$. Go round the infinite circle, coming up the negative reactive axis, until the arc 0.375 is met.

$$z_{in.} = 0 - j1$$

$$Z_{in.} = -j300 \text{ ohms (capacitive).}$$

Example.—A transmission line of characteristic impedance 300 ohms is 0.15λ long. The input impedance is resistive and equal to 900 ohms. What is the impedance at the end of the line?

$$z_{in.} = \frac{900}{300} = 3. \text{ Find the point } (3, 0).$$

Note its n value, $n'' = 0.25$. Travel round *anticlockwise* corresponding to a length of line 0.125λ, i.e. subtract 0.15 from 0.25 obtaining $n' = 0.10$.

$$z_{ter} = 0.5 + j0.6$$

$$Z_{ter} = 150 + j180 \text{ ohms.}$$

Example.—A transmission line of characteristic impedance 300 ohms is short-circuited at its far end. The input impedance is +j600 ohms. What is the length of the line?

$$z_{ter} = 0 + j0$$

$$z_{in.} = \frac{0 + j600}{300} = 0 + j2.$$

Find the point (0, 0). Travel round the infinite circle until the point (0, 2) is reached; $n = 0.177\lambda$.

34. Deduction of terminating impedance from standing wave measurements

Let the voltage standing-wave ratio be s and let it be expressed as the ratio of maximum to minimum, i.e. a number greater than unity. At a point on the line where there is a voltage maximum the line appears to be purely resistive—the resistance being sZ_0 ohms. At a voltage minimum the line is also resistive and equal to Z_0/s ohms. From these facts and knowing the distance of a maximum or a minimum from the termination one can deduce the terminating impedance.

Example.—The voltage standing wave ratio on a line is 3:1. The voltage maximum is distant 0.4λ from the termination. Find the terminating impedance.

$$\text{Impedance at voltage maximum} = 3Z_0$$

$$\text{Normalised impedance at maximum} = 3.$$

Find the point (3, 0). Note its n value, $n'' = 0.25$. Travel round *anticlockwise* a distance 0.4λ, i.e. subtract 0.4 from 0.25. This cannot be done, so subtract from 0.5 + 0.25, i.e. from 0.75. One finds $n' = 0.75 - 0.4 = 0.35$.

Termination is $(.8 - j1)Z_0$ ohms.

35. Use of admittances

Impedances are not really convenient in transmission line work. It is unusual, for

example, to connect a resistance and condenser in series across a transmission line, forming an impedance $r + jx$. Commonly one connects elements in shunt across the line. When working with shunt elements one uses admittances. The reciprocal of resistance is “conductance”, of reactance is “susceptance”, and of impedance is “admittance”. Let R and jX be a resistance and reactance connected in shunt (fig. 29). Their impedance Z is given by

$$\frac{1}{Z} = \frac{1}{R} + \frac{1}{jX} = \frac{1}{R} - \frac{j}{X}$$

Writing $Y = \frac{1}{Z}$, $G = \frac{1}{R}$ and $B = -\frac{j}{X}$ this becomes:

$$Y = G + jB,$$

so that shunt elements can be added if expressed as conductance and susceptances. Note that an inductive element, which is a positive reactance, is a negative susceptance. Similarly a capacitive element is a negative reactance but a positive susceptance. The transmission line calculator can be used for normalised admittances in the same way as for normalised impedances.

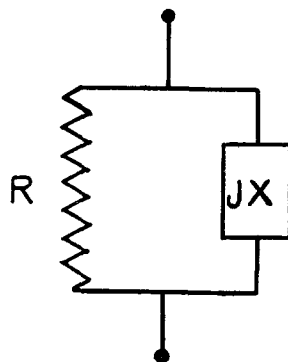


Fig. 29.—Shunt elements

36. Unfortunately the reactive axis on the cartesian circle diagram (fig. 185, A.P.1093-E) has been marked “inductive” and “capacitive”. It would have been better to have called the upper half “positive reactance” and the lower “negative reactance”. In using the diagram for admittances the upper half denotes “positive susceptance” and the lower half “negative susceptance”.

Elimination of standing wave on line by stub (Stub-matching)

37. If a line is terminated by a normalised impedance (or admittance) of unity, the power is all absorbed at the termination and there is no standing wave between the generator and the termination; the standing wave ratio is

unity. The method of eliminating a standing wave by an inductive stub is indicated in the following example:—

Example.—Show how to eliminate a voltage standing wave ratio of 3:1 on a transmission line.

At a voltage maximum on the line the line looks resistive—the normalised resistance being 3, since this is the standing wave ratio. Hence, at a voltage maximum, the normalised conductance is $1/3$. Find the point $(1/3, 0)$ on the circle diagram. Travel round the circle until the conductance is unity. The n value is $n = 0.168$. The susceptance is $+j1.15$. Thus if we move along the line from a voltage maximum, towards the generator, a distance 0.168λ the line will have a normalised admittance $1 + j1.15$ at that point. If the susceptance $j1.15$ can be cancelled out, by adding in shunt a susceptance $-j1.15$, the admittance will then be unity and there will be no standing wave between that point and the generator. Let the required susceptance be supplied by a short-circuited stub applied across the line at this point. Let the stub have the same characteristic

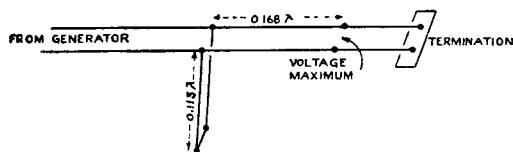


Fig. 30.—Stub match for 3:1 standing wave

admittance or impedance as the main line. A short circuit is zero impedance or infinite admittance. Hence find the point ∞ on the diagram. Its n value is $n' = 0.25$. Travel clockwise round the “infinite circle” coming up the vertical axis until the point $(0, -1.15)$ is reached. Note the n value, $n'' = 0.363$. The required length of stub is

$$n'' - n' = 0.363 - 0.25 \\ = 0.113\lambda.$$

The arrangement for matching out the 3:1 standing wave is therefore as shown in fig. 30.

Line with loss

38. **Example.**—A transmission line of characteristic impedance 80 ohms is terminated in a resistance of 240 ohms. The length of the line is 8.6 wavelengths and the loss of the line is 2dB. Find the input impedance.

$z_{\text{termin}} = 3 + j0$. Find the point $(3, 0)$.

$$n' = 0.25$$

Subtract $\cdot 1$

$n'' = 0.35$. Travel round until arc 0.35 is met. This circle is marked 3dB. Add 2 and obtain 5dB. Travel in until 5dB circle is met.

Answer is:—

$$Z_{\text{input}} = 1.0 - j0.7$$

$$Z_{\text{input}} = 300 - j210 \text{ ohms.}$$

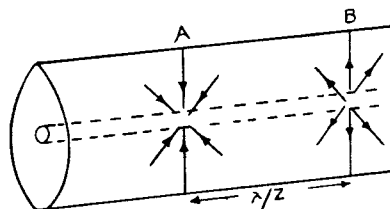


Fig. 31.—Coaxial feeder

R.F. CABLES

39. Up to about 1,000 Mc/s the attenuation in polythene cables is mainly due to losses in the metal (“copper” losses) and the attenuation varies as the square root of the frequency. Beyond 1,000 Mc/s the dielectric loss begins to become appreciable and at 3,000 Mc/s the loss is about half dielectric and half “copper”. The dielectric loss varies directly as the frequency and at very high frequencies (10,000 Mc/s or more) it is the important factor.

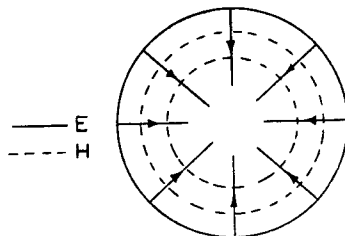


Fig. 32.—Fields with inner removed

Loss in a coaxial and transition to waveguides

40. In an air-spaced coaxial line, the main copper loss is on the inner conductor. At high frequencies, the current tends to travel in the surface of the metal (skin effect) and the

amount of metal available is therefore much smaller on the inner conductor than in the outer conductor. In addition, the inner conductor is troublesome, since it requires supports and these upset the transmission and may give rise to sparking. The efficiency of transmission can be much improved, therefore, by removing the inner conductor. This gives a single hollow pipe or waveguide. When the inner conductor is taken out the fields are altered. The diagram fig. 31 shows, roughly, the electric field at two points A and B, half a wavelength apart. On removal of the inner conductor, the cross section at A is as shown in fig. 32. The electric lines are left "hanging in the air". They cannot join across since they are in opposing directions. The only way they can complete their path is by turning round and extending down the pipe so as to join with the electric lines at B, which are also incomplete but pointing in the opposite direction to the lines at A. The result is shown in fig. 33. The magnetic lines in the coaxial form circles as shown in fig. 32. The removal of the inner conductor merely causes a slight change in the distribution of magnetic field, there being no need to connect up with the field at other parts of the pipe. The wave in the pipe differs notably from the wave on a transmission line in that the electric field has both a transverse and a longitudinal component; the magnetic field has only a transverse component. Seizing on the fact that E has a longitudinal component the wave is referred to as an "E wave".

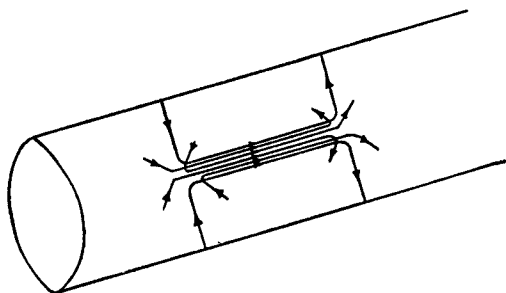


Fig. 33.—Circular waveguide

APPROACH TO RECTANGULAR GUIDE

41. The twin wire transmission line gives rise to difficulties at high frequencies due to skin effect on the wires, in the same way as the inner conductor of the coaxial causes loss. By using metal strips instead of wires—i.e. a twin strip line—the loss is considerably reduced since there is now a large amount of surface

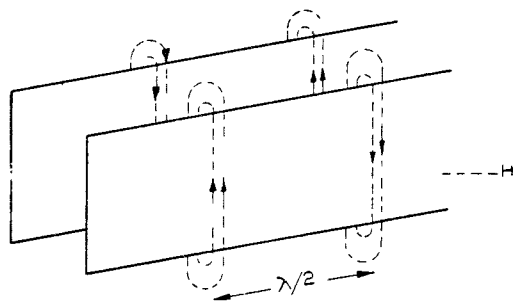


Fig. 34.—Magnetic fields on strip lines

skin in which the current travels. However, the twin strip line (see fig. 34) suffers from the disadvantage that it is not a screened line. Thus in transmission it is liable to radiate, especially at junctions, bends and such discontinuities. Similarly, in reception, it is liable to pick up interference. The magnetic lines forming closed loops outside the strips are shown in fig. 34. In order to enclose the field completely between the strips, metal plates must be added at the top and bottom. The oscillating magnetic field is unable to penetrate these plates. It must, however, still form closed loops. This is done by the magnetic field, at one cross-section, bending over at the top and bottom and joining up with the magnetic field approximately $\lambda/2$ away (see fig. 35). The electric field is less disturbed by the addition of the top and bottom plates. It is merely caused to concentrate more towards the centre of the cross-section and be zero at the top and bottom so that the condition $E_{\text{tangential}} = 0$ is satisfied. In fig. 36 there is shown a cross section. This wave in the rectangular pipe is a "waveguide mode". It is an 'H' wave since there is a component of magnetic field in the direction of propagation.

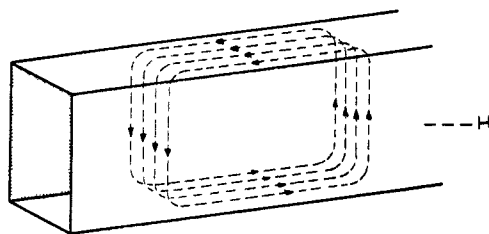


Fig. 35.—Rectangular guide

Synthesis of H wave in rectangular guide

42. The previous qualitative approach does not give sufficient information adequately to explain some fundamental points on waveguides. A simple analysis can be made by the method of Brillouin. Take two equal plane

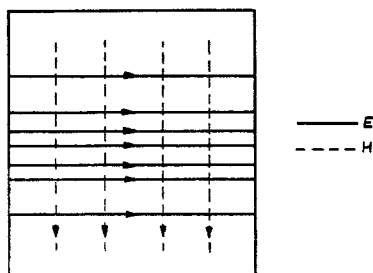


Fig. 36.—Cross section of guide

waves travelling with their directions of propagation parallel to the plane of the paper but in different directions in that plane (see fig. 37). Assume H to be parallel to plane of paper and E perpendicular to plane of paper. The full lines are wavefronts on which H (and E) is a maximum or a minimum, while the dotted lines are wavefronts on which H (and E) are zero. This is indicated by arrows at the ends of the wavefronts. When the two waves are combined, by adding vectorially, the magnetic field forms closed loops characteristic

of the pattern already deduced for a rectangular guide. The result of adding the electric fields in the two waves are not shown in fig. 37. However, they are fairly simple. The electric field is zero at all points along lines such as PQ, RS, TV. Half-way between PQ, RS, TV, the electric field takes its maximum values at the points where the magnetic field has its maximum value.

43. Provided that the boundary conditions are satisfied, metal plates can be inserted so as to separate off desired portions of the field pattern in fig. 37. Thus, two metal plates RS and TV can be inserted separating off the magnetic loops. The magnetic field is wholly tangential at these metal plates and the electric field at the plates vanishes. The wavefronts in fig. 37 extend indefinitely perpendicular to the paper. To form a waveguide from the portion enclosed between RS and TV, the side walls have to be added as shown in fig. 38. There is no difficulty in inserting these side walls since the magnetic field is tangential and electric field normal. The spacing, a , between the side walls can be chosen to be any

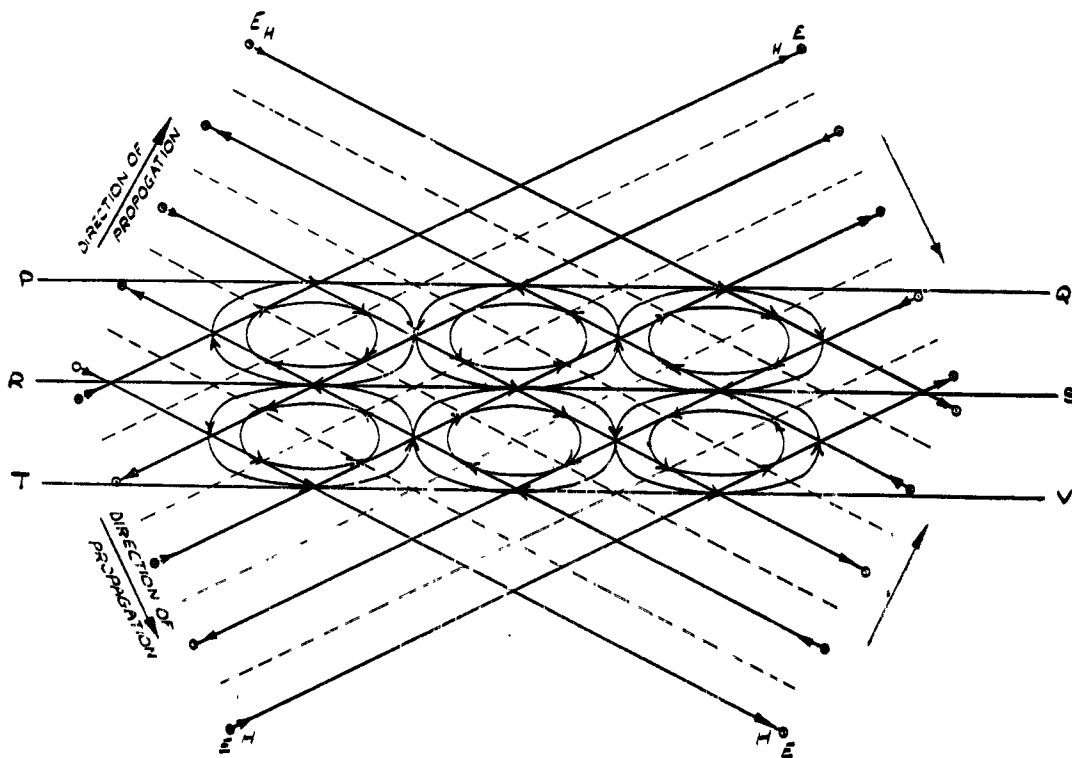


Fig. 37.—Synthesis of H_0 wave

value whereas the top and bottom plates must be exactly in the correct positions to fit the pattern. The wave shown in fig. 38 is an H_{01} mode. The suffixes 0 and 1 refer to the number of cycles of variation in, say, the electric field in moving along the x and y directions (fig. 38). An H_{02} wave would be obtained by enclosing two rows of magnetic loops in fig. 37 by planes TV and PQ giving the pattern shown as a cross-section in fig. 39.

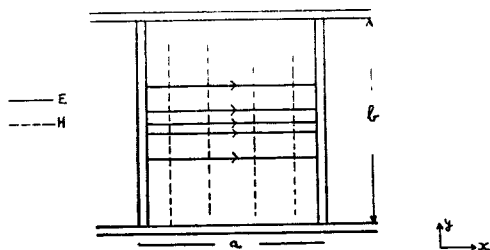


Fig. 38.—Fitting side walls

Relation between wavelength in space and in the guide

44. Referring back to fig. 37, the original plane waves are shown moving down to the right and up to the right. The waveguide pattern, enclosed between RS and TV, therefore moves to the right and can be regarded as a type of wave, i.e. an 'H' wave as explained

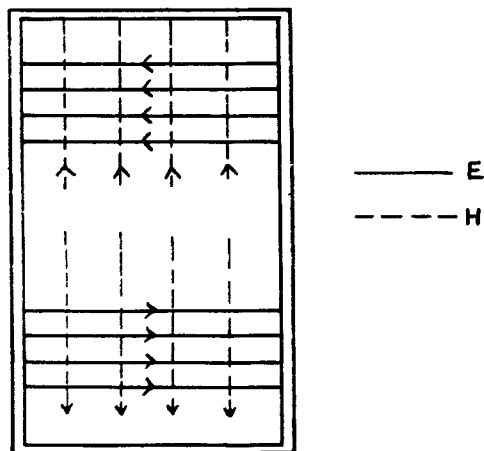


Fig. 39.—Guide with H_{02} mode

previously. The loops of magnetic field are marked alternately with anticlockwise and clockwise directions so that a "wavelength" is two loops and each loop is $\lambda_g/2$ where λ_g means "wavelength in the guide". The wavelength of the radiation in free space (wavelength in "air") is λ_a and the distance between two

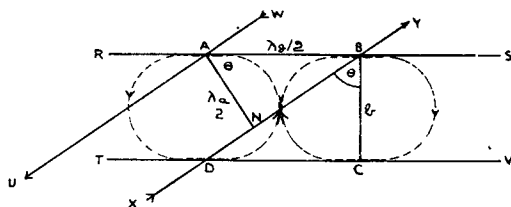


Fig. 40.—Relationship between λ_g and λ_a

of the full-line wavefronts in fig. 37 is $\lambda_a/2$. A portion of fig. 37 is shown separately in fig. 40. UW and XY are two wavefronts $\lambda_a/2$ apart. AB is the distance between the centres of two magnetic loops and is $\lambda_g/2$. BC is the height of the guide or distance between the top and bottom plates RS, TV and is denoted by b. Let angle CBD = θ and let AN be perpendicular to BD.

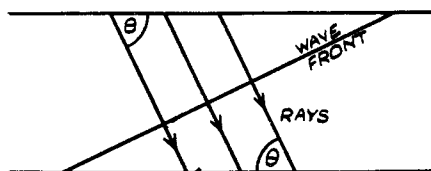


Fig. 41.—Rays of downgoing wave

In triangle ABN we have $\cos \theta = AN/AB$
 $= \lambda_a/\lambda_g$.

In triangle BCD we have $\tan \theta = CD/BC$
 $= \lambda_g/(2b)$.

Hence, $\sin \theta = \tan \theta \times \cos \theta$
 $= \lambda_a/(2b)$.

Now $\sin^2 \theta + \cos^2 \theta = 1$.

Hence $\lambda_a^2/(2b)^2 + \lambda_a^2/\lambda_g^2 = 1$
 i.e. $1/(2b)^2 + 1/\lambda_g^2 = 1/\lambda_a^2$
 or $1/\lambda_g^2 = 1/\lambda_a^2 - 1/(2b)^2$.

This equation enables λ_g to be calculated when λ_a and b are known.

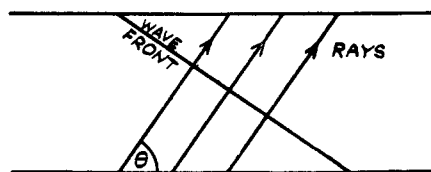


Fig. 42.—Rays of upgoing wave

45. The interpretation of the angle θ is also interesting. Previously in this treatment we have been concerned with the wavefronts of the two component waves. We may also

consider these waves from the ray point of view. A ray is at right angles to a wavefront and rays can be drawn all along the wavefront as desired. In fig. 41 rays are shown as arising from the downward travelling wave and the angle θ (from fig. 40) is as shown. Similarly for the upward travelling wave we obtain fig. 42. Combining appropriate rays we obtain the effect of a series of rays striking the upper and lower plates at a glancing angle θ and being reflected each time they strike the plates (fig. 43). This is the "bouncing wave" sometimes referred to and gives the interpretation of the angle θ in the equations.

Example.—Find θ and λ_g for a waveguide measuring $2\frac{1}{2}$ in. $\times$ 1 in. when the free space wavelength is 9.1 cms.

We have $\sin \theta = \lambda_a/(2b)$.

The dimension b is $2\frac{1}{2}$ in. and $2b = 5$ in. = 5×2.54 cm.

$$\begin{aligned}\text{Hence, } \sin \theta &= 9.1/(5 \times 2.54) \\ &= 1/1.4 \\ &= 1/\sqrt{2} \\ \therefore \theta &= 45^\circ.\end{aligned}$$

$$\begin{aligned}\text{Again, } \tan \theta &= \lambda_g/(2b) \\ \tan 45^\circ &= 1 \\ \therefore \theta &= 2b \\ &= 5 \text{ in.} \\ &= 12.7 \text{ cm.}\end{aligned}$$

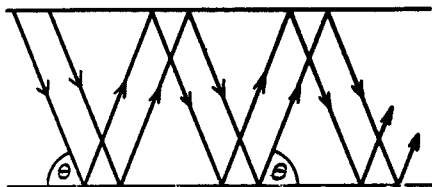


Fig. 43.—Bouncing wave effect

46. This is the guide used in airborne 'S' band equipment. A bouncing angle of 45 deg. is a convenient value to assume if one wishes to design a waveguide to suit a given free space wavelength. Note that the narrow dimension 1 in. does not enter in the calculation—as already noticed, the distance apart of the side plates is not critical.

Example.—Find λ_g for a 3 in. $\times$ 1 in. guide at $\lambda_a = 10$ cms. We do not need to calculate θ and the formula

$$\frac{1}{\lambda_g^2} = \frac{1}{\lambda_a^2} - \frac{1}{(2b)^2}$$

may be used. Thus,

$$\begin{aligned}\frac{1}{\lambda_g^2} &= \frac{1}{10^2} - \frac{1}{(6 \times 2.54)^2} \\ &= \frac{15.24^2 - 10^2}{(10 \times 15.24)^2} \\ \lambda_g &= \frac{10 \times 15.24}{5.24 \times 25.24} \\ &= 13.2 \text{ cm.}\end{aligned}$$

Example.—1 in. $\times$ $\frac{1}{2}$ in. guide at $\lambda_a = 3.2$ cm. To be done as an exercise.

Cut-off in a rectangular waveguide

47. The formulæ relating glancing angle θ with λ_a and b and guide wavelength, λ_g , with λ_a and b are:—

$$\sin \theta = \lambda_a/(2b)$$

$$\frac{1}{\lambda_g^2} = \frac{1}{\lambda_a^2} - \frac{1}{(2b)^2}$$

for an H_{01} wave.

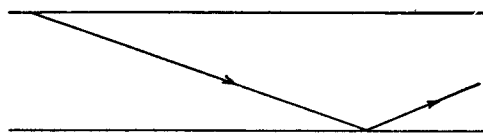


Fig. 44.—Propagation well below cut-off

If λ_a is very small compared with $2b$, then θ is small and the rays do not bounce often (fig. 44). At the same time $1/\lambda_a^2$ is large compared with $1/(2b)^2$ and λ_g is approximately equal to λ_a . The wave almost behaves like a wave in free space.

If $\lambda_a = 2b/\sqrt{2}$, then $\sin \theta = 1/\sqrt{2}$, $\theta = 45$ deg. and $\lambda_g = 2b = \sqrt{2}\lambda_g$. This is the usual sort of condition for a waveguide in practice. The "wavelength" in the guide is about 40 per cent. greater than the wavelength in free space.



Fig. 45.—Propagation at cut-off

If $\lambda_a = 2b$, then $\sin \theta = 1$, $\theta = 90$ deg. $\lambda_g = \infty$. In this case the rays bounce normal to the top and bottom plates (see fig. 45) and no power is conveyed down the pipe.

48. If $\lambda_a > 2b$, $\sin \theta > 1$ and λ_g is imaginary. The wave cannot exist in the form of a travelling wave moving down the pipe. The waveguide thus acts like a high pass filter; only high frequencies or low wavelengths are transmitted. The change from transmission to no transmission occurs as the free space wavelength passes through the value $2b$. The quantity $2b$ is thus called the critical or cut-off wavelength, since wavelengths higher than this will not propagate. The cut-off wavelength is called λ_c and we have

$$\lambda_c = 2b.$$

The equation for λ_g can then be written

$$\frac{1}{\lambda_g^2} = \frac{1}{\lambda_a^2} - \frac{1}{\lambda_c^2}$$

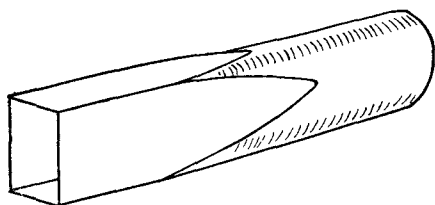


Fig. 46.—Distortion from rectangular to circular guide

49. In this form it is applicable to any waveguide for which λ_c is known. Even when λ_a is greater than λ_c some energy goes into the guide and falls off exponentially with distance. This type of propagation above cut-off wavelength is called an evanescent wave or an attenuated wave. A detailed investigation shows that there is no real flow of energy in an evanescent wave. It is in fact "reactive", magnetic field goes into the pipe and then returns to the generator; a quarter of a cycle later electric field is going in and then returning and so on. The amplitude of these reactive fields falls off exponentially with distance from the generator.

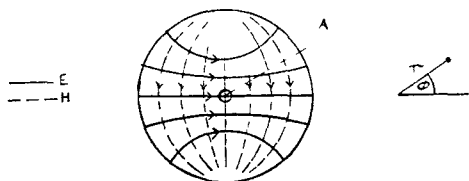


Fig. 47.— H_{11} wave in circular pipe

CIRCULAR WAVEGUIDES

H_{11} wave in circular pipe

50. Beginning with a rectangular guide carrying an H_{01} wave, let the rectangular pipe be gradually distorted into a circular pipe (fig. 46). The wave pattern in the cross-section is also distorted, the electric lines bending so as to finish always perpendicular to the walls, and the magnetic lines bulging out so to lie parallel to the walls. Thus, the pattern of figs. 36 or 38 becomes that of fig. 47. In designating this H wave in the circular pipe it is usual to refer the variations of field to polar co-ordinates (φ , r). Taking first the angular variation consider a rod OA pivoted at O and moved round (see fig. 47). Consider, for example, the radial component of electric field. In the position A_1 (see fig. 48) this field

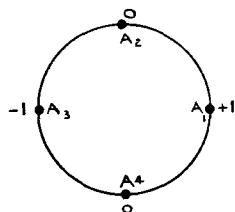


Fig. 48.—Azimuthal variation H_{11} mode

is strong and directed radially outwards; at A_2 it is zero; at A_3 strong and radially inwards; at A_4 zero again; and so on. The field thus varies like $\cos \varphi$ which is $+1, 0, -1, 0$ when $\varphi = 0$ deg., 90 deg., 180 deg., 270 deg. respectively. Taking the general expression to be $\cos n\varphi$ it is seen that in our case $n = 1$. The first suffix in the designation is therefore 1. For the second suffix we look at the variation along a radius. Along the radius to A_1 , the electric field is strong at the centre and weakens to the edge. This is regarded as one "variation" of field and the second suffix is 1. The wave is referred to as an H_{11} . It is, of course, necessary to mention the fact that it is in a circular guide so that the meaning of the suffixes can be understood. The cut-off wavelength for an H_{01} wave in a rectangular guide is twice its height. So we would expect that the cut-off wavelength for an H_{11} wave in a circular guide would be about twice its diameter or four times its radius. Accurate investigation gives the value

$$\lambda_c = 3.42 \times \text{radius}.$$

E_{01} wave in circular pipe

51. This is the wave obtained by removing the inner conductor from a coaxial feeder,

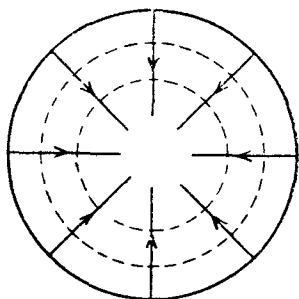


Fig. 49.— E_{01} wave

and previously discussed. The field at a cross-section is reproduced in fig. 49. There is no angular variation in the field so that $n = 0$ in the expression $\cos n\phi$. There is one radial variation. Hence it is an E_{01} . Its cut-off wavelength is given by $\lambda_c = 2.61 \times \text{radius}$. The H_{11} wave has the highest cut-off wavelength of all the circular modes. The E_{01} comes next. If r were the radius of the pipe and a free space wavelength of $3r$ were sent into the pipe then an H_{11} wave would propagate

but an E_{01} would be evanescent. Sometimes circular guide with a radius equal to $\lambda_a/3$ is used to convey power in an H_{11} wave instead of rectangular guide with an H_{01} .

52. However, circular guide with H_{11} is never used except in short lengths. Rectangular guide is preferred. Referring to fig. 47, the electric field is shown horizontally polarised. But any slight ellipticity or curve or other distortion would easily tend to slip the polarisation off the horizontal. There is nothing in a circular guide to hold the plane of polarisation. In a rectangular guide the polarisation is held by the wide sides and since we usually wish to know accurately the polarisation of the wave at the end of the guide, rectangular—in spite of being more expensive than circular—is normally used.

CURRENTS IN WAVEGUIDES

Currents in walls of rectangular guide carrying H_{01} wave

53. The guide has been investigated from the field point of view but it is important to know

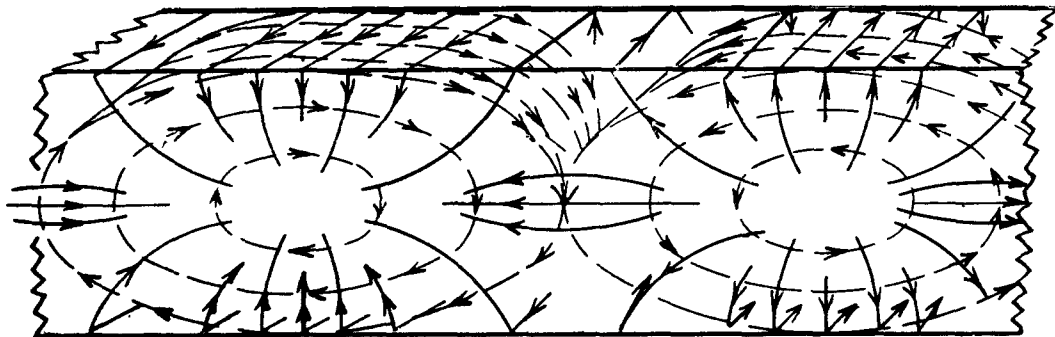


Fig. 50.—Currents in walls of guide

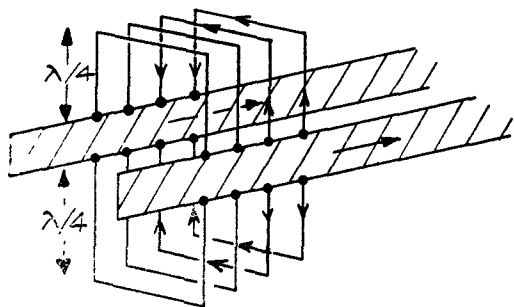


Fig. 51.—Transmission line with stubs

how the currents flow on the inside of the walls of a rectangular pipe. The currents are easily deduced by remembering that magnetic

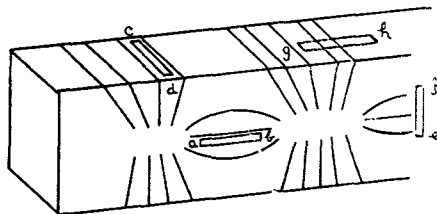


Fig. 52.—Slots in walls of guide

field lies parallel to the walls and the current at any point on the walls is perpendicular to the magnetic field there. The magnetic field at the wall is shown by dotted lines in fig. 50 and the lines of current flow have been sketched in. The whole pattern of current lines moves up the pipe with the wave.

54. The picture shown in fig. 50 suggests another way of considering a waveguide. There are essentially two current flows; first *longitudinal*, along a region on either side of the centre line of the wide side; and second *transverse*, across the edge of the wide side and running over the narrow side. The same effect is obtained by having two parallel metal strips acting as a transmission line and a set of shorted quarter wave stubs above and below (see fig. 51).

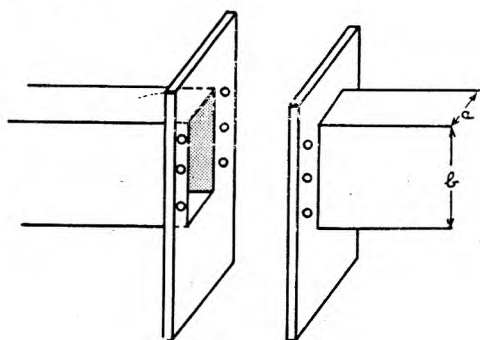


Fig. 53.—Flanges

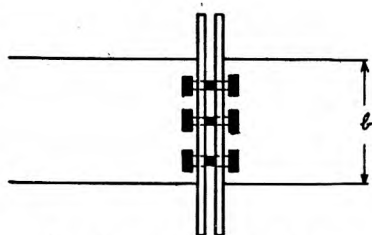


Fig. 54.—Bolting of flanges

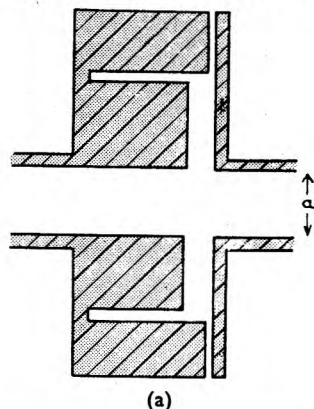
Slots and joints in rectangular guide with H_{01} wave

55. We consider again the currents on the inside walls of the rectangular guide, redrawn in fig. 52. If a thin slot *ab* is cut (for example, by a milling machine) along the centre of the wide side, the current pattern is not disturbed and the slot has no effect on the current pattern or on the wave inside the guide. Such a slot is useful in enabling one to insert a probe into the guide and measure the field and deduce whether any standing waves are present. Another slot which may be cut without up-

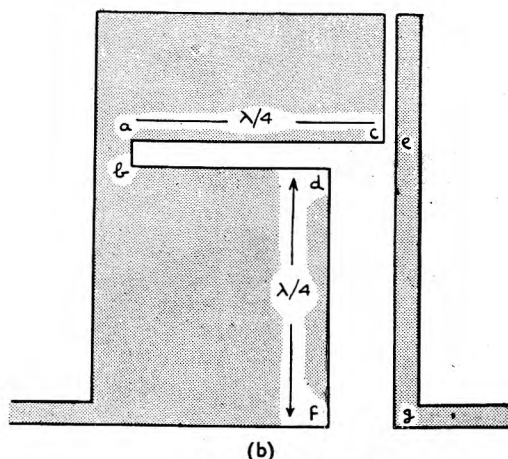
setting the wave is *cd*, across the narrow side. On the other hand, slots such as *ef* or *gh* will completely stop the flow of current and will upset the wave badly. Charge will accumulate at the edges of such slots and they will radiate power out into space.

56. Suppose it is desired to join together two pieces of guide. Flanges are soldered on to the guides as shown in fig. 53 and the guides bolted together as shown in fig. 54. It is clear that care must be taken to tighten up the bolts properly, especially the middle one. A gap between the flanges on the long dimension would be a radiating slot of type *ef* in fig. 52. On the other hand, a gap across the top and bottom narrow dimensions would be of the non-radiating type *cd* in fig. 52 and would not matter. Indeed as shown in fig. 53 and 54 bolts are not required at the top and bottom—at least, not electrically.

57. Sometimes, instead of relying simply on bolts a *choke joint* is used. This is illustrated in fig. 55(a) and partly again on a larger scale in fig. 55(b). Regard the “ditch” *abcd* as a



(a)



(b)

Fig. 55.—Choke joint

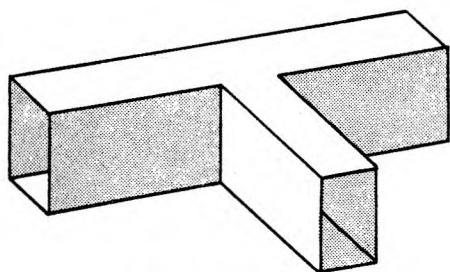


Fig. 56.—Series arm

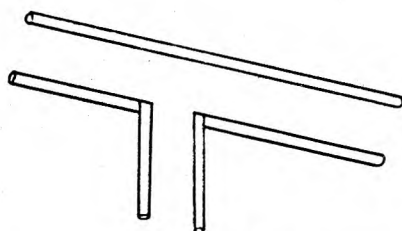


Fig. 57.—Analogue of series arm

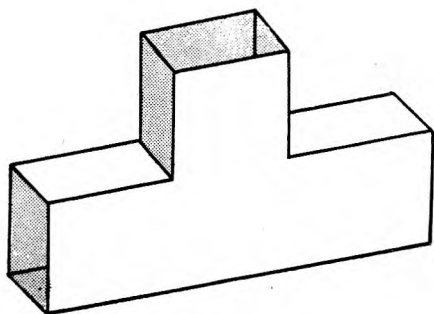


Fig. 58.—Shunt arm

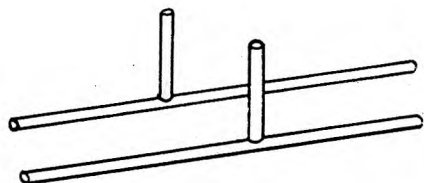


Fig. 59.—Analogue of shunt arm

transmission line $\lambda/4$ long. It is short circuited at ab so there is an infinite impedance between c and d . Between c and e there is an indefinite impedance depending on the goodness of the contact between c and e . However, this indefinite impedance comes in series with the infinite impedance and we obtain in any case an infinite impedance between d and e . This transforms through $\lambda/4$ line to a short circuit between f and g . The wall of the guide is therefore effectively short-circuited.

Series and shunt side arms

58. It is easily deduced from the current picture in fig. 52 that the junction in fig. 56 is a series junction and the equivalent transmission line circuit is as in fig. 57. Similarly fig. 58 shows a shunt junction equivalent to fig. 59.

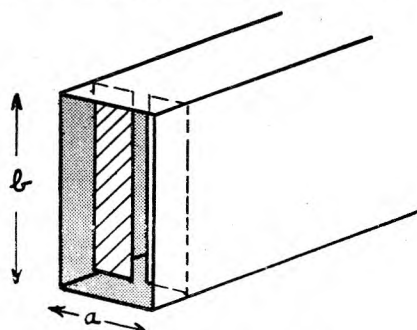


Fig. 60.—Capacitive iris



Fig. 61.—Field near iris

WAVEGUIDE TECHNIQUE

Irises in rectangular guide with H_{01} wave

59. The previous paragraph suggests that we might use series or shunt waveguide stubs for eliminating standing waves in a guide in the same way as for transmission lines. However, a simpler method is provided by the use of irises or thin diaphragms which are inserted at an appropriate place in the guide. Consider the iris shown in fig. 60 and again, in plan drawing, in fig. 61. Note how in fig. 61 there is more than the normal number of lines of electric force in the region round the iris due

to extra lines leaving and arriving at its surface. There is thus additional electric energy or capacity concentrated there and the iris acts like a condenser. The equivalent transmission line or circuit picture is shown in fig. 62.

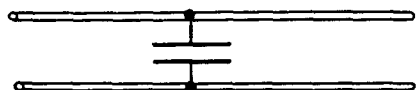


Fig. 62.—Analogue of capacitive iris

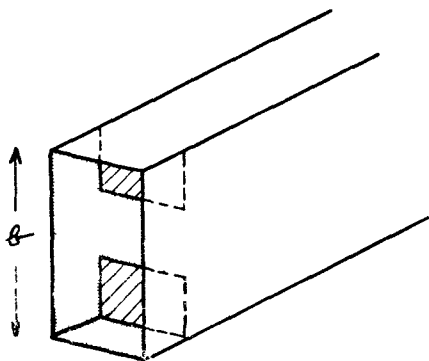


Fig. 63.—Inductive iris

60. Again consider the iris in fig. 63. It can be seen that this iris will not greatly disturb the distribution of electric field which is mainly concentrated near the middle of the guide. The magnetic loops in the guide will, however, be distorted by the iris so as to form an additional concentration of magnetic energy so that this iris acts like a coil and is equivalent in transmission lines and circuits to fig. 64.

61. The capacitive iris is not much used in practice owing to its tendency to spark over under high power and the inductive iris is preferred.

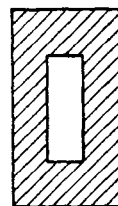


Fig. 64.—Analogue of inductive iris

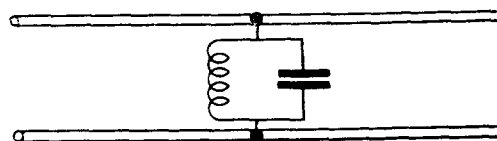
62. By putting an inductive and capacitive iris together we obtain a resonant window which is equivalent to a parallel tuned circuit (see fig. 65(a) and 65(b)).

63. Various other types of iris, etc. will be found in practice and it is easy to see that any non-lossy obstruction or obstacle in a waveguide will, in general, act as a reactance. For example, a rectangular piece of distrene is

sometimes used. By altering its orientation in the guide different reactive effects are obtained and by sliding it along the guide the point of concentration of the reactance may be varied. For low power purposes a metal screw passing through the centre of the wide face provides a convenient variable reactance (see fig. 66). It is capacitive if its length inside the guide is less than $\lambda/4$.



(a)



(b)

Fig. 65.—Resonant iris

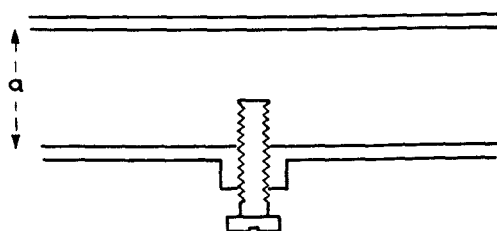


Fig. 66.—Captive screw

Rotating joint using H_{01} to E_{01} transformer

64. A rotating joint in a waveguide run must clearly employ circular waveguide and in most applications an E_{01} wave is used in the circular guide. To generate this wave the rectangular guide is brought in at right angles to the circular guide (see fig. 67). As shown in fig. 68, the electric field, of the H_{01} wave in the rectangular guide, expands into the circular guide and creates some of the characteristic half-loops of electric field for the E_{01} wave in the circular guide. After travelling a little way up the guide, the E_{01} is fully developed. A choke type of joint provides a convenient method of obtaining a rotating joint in the circular pipe and the wave may then be launched back into a rectangular guide as shown in fig. 69. The bottom half of the structure might be connected

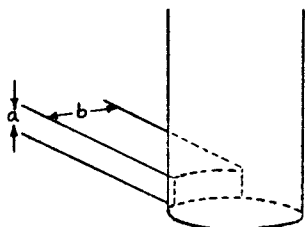


Fig. 67.— H_{01} to E_{01} transformer

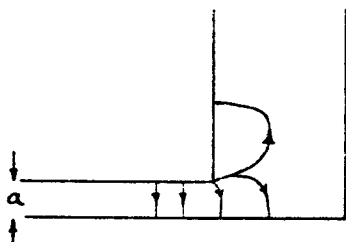


Fig. 68.—Field in transformer

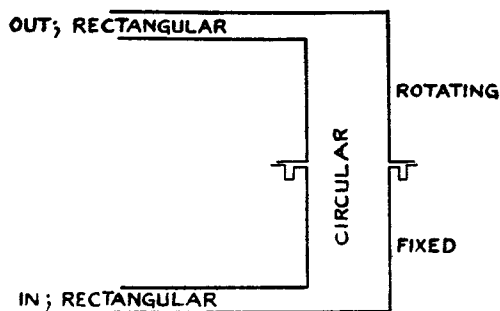


Fig. 69.—Rotating joint

to a transmitter and the top half to a rotating mirror or scanner. It is important that the same amount of power should emerge from the mirror in all positions. This is ensured by using the symmetrical E_{01} wave (fig. 70) which "looks the same", to the output rectangular guide, from whatever orientation (fig. 71). Unfortunately the E_{01} mode has a lower cut-off wavelength than the H_{11} mode in circular guides. Hence, some H_{11} is very likely to be present in the circular pipe. This mode is polarised as shown in fig. 72. Consequently, the circular guide "looks different", to the output rectangular guide, in different orientations, and for example, in position (1) of fig. 71 no output would be obtained at all with the H_{11} mode.

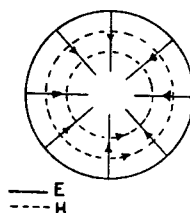


Fig. 70.—Field in circular guide with E_{01} mode

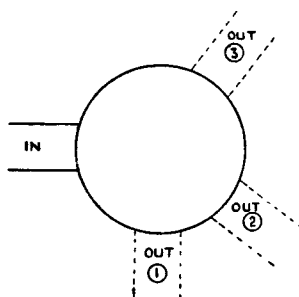


Fig. 71.—Variation of output plane

65. The H_{11} mode is filtered out by a metal ring filter, placed centrally in the circular guide (see fig. 73). The electric lines of force in the E_{01} mode (see fig. 70) are perpendicular to the ring, the boundary conditions of the wave are thus satisfied and the wave moves through the ring without interruption. But the electric lines of force in the H_{11} mode

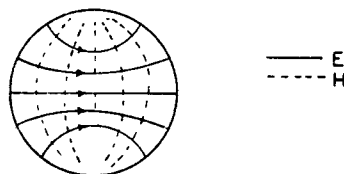


Fig. 72.— H_{11} field

(see fig. 72) are not, in general, perpendicular to the metal ring. A reflected wave is called into play of equal amplitude to the incident wave and of opposite electric phase, so that the resultant electric field vanishes at the ring. The ring thus reflects back the H_{11} wave. In practice the ring is supported by a thin wafer of good insulating material as, for example, polythene.

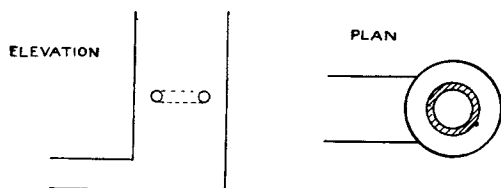


Fig. 73.—Ring filter

Bends

66. If a waveguide is bent round too sharply, part of the wave is reflected back due to the discontinuity. The minimum radius of curvature which can be allowed is about λ_g . When it is necessary, owing to space considerations, to bend sharply, a right-angle bend with cut-off corner is employed (see fig. 74). This bend can be in either the narrow dimension or wide dimension of the guide. The distance d has to be determined by the designers for each particular guide size, but it is not very sensitive to frequency changes. When correctly designed, the cut-off corner gives rise to negligible reflected wave.

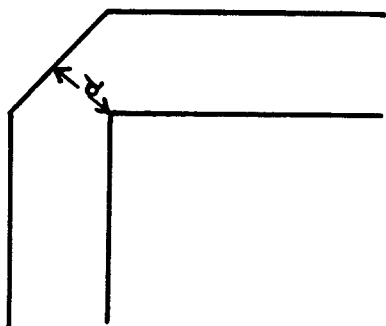


Fig. 74.—Right angle bend

Launching waves in guides

67. The most common case is the H_{01} wave in a rectangular guide which has to be launched from a magnetron. The magnetron has its output in the form of a coaxial line and this is brought in at right angles to the guide with the inner conductor running about $\lambda/4$ inside the guide (fig. 75). This then acts like a little aerial and generates waves whose electric vector is perpendicular to the wide walls of the guide. The field distribution set up inside the guide by the probe will be very complicated. It can be represented, so to speak, by a kind of Fourier series, i.e. by the sum of a large number of waveguide modes, of various amplitudes and phases. Outstanding amongst these modes will be the H_{01} mode since its electric field is perpendicular to the wide walls and fairly

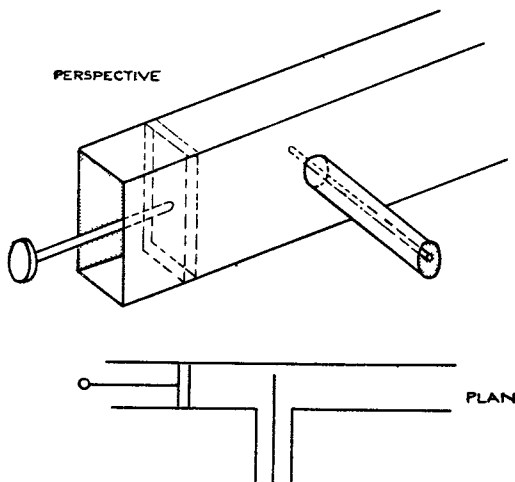


Fig. 75.—Launching of H_{01} wave

evenly distributed across the centre of the guide. All the modes will attempt to propagate up the guide, but by making the guide sufficiently small the higher modes will be evanescent and will damp out leaving only the H_{01} . Generally speaking, the idea is to generate a field of which the required mode is a predominant part, and to attempt to eliminate unwanted modes by adjusting guide dimensions or by using filters.

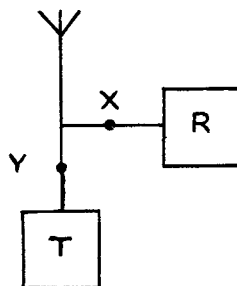


Fig. 76.—Common T and R aerial

68. Referring to fig. 75 a piston is provided for matching purposes. If the piston is omitted, the guide is closed $\lambda_g/4$ from the probe. At high powers the probe sparks over and has to be covered with polythene or terminated in a large sphere; the wide dimension of the guide is also usually increased to help to stop spark-over.

COMMON T AND R

Use of common T and R

69. In the pulse method of radiolocation (which is the method normally used) no

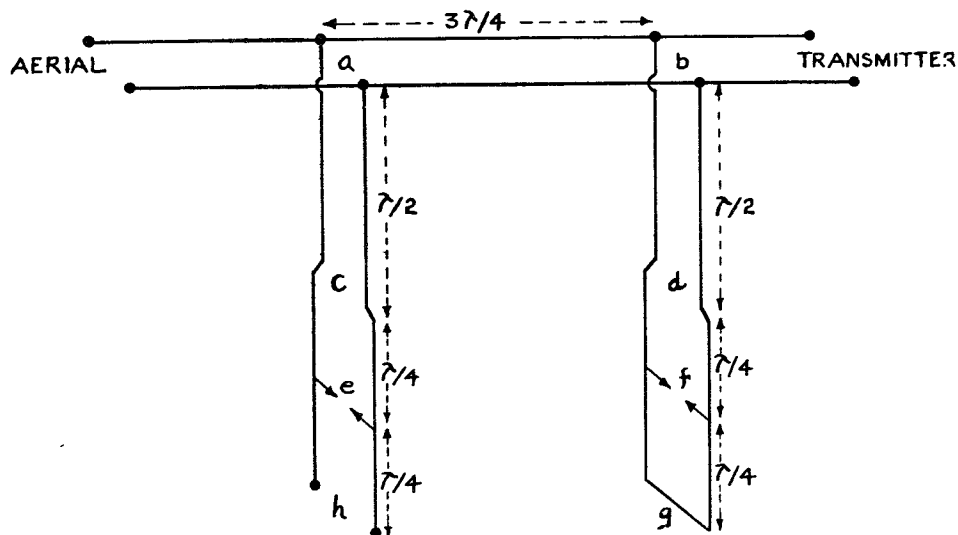


Fig. 77.— $1\frac{1}{2}$ metre common T and R switching

received signals are expected to be detected when the transmitter pulse is being radiated and there is no transmission during the reception of echoes. The same aerial can therefore be used for transmission and reception; this is termed *common T and R*. However, owing to the high peak power of the transmitter, the receiver would be damaged by the transmitter pulse if the simple connections of fig. 76 were used. It is necessary to introduce a device at X which will short circuit the line to earth or otherwise during transmission, and prevent the transmitted pulse entering the receiver. It is usual also to have a device at Y which will prevent the received signal entering the transmitter, but this is not so essential since the transmitter output circuit will not absorb much received power and the device at Y could be omitted in a "poor man's radar".

$1\frac{1}{2}$ metre common T and R

70. The arrangement shown in fig. 77 is typical for a ground station. Airborne sets use the same principles but differ in detail. Spark gaps are provided at e and f. These are enclosed in glass envelopes with argon and mercury vapour at $\frac{1}{2}$ atmosphere pressure.

71. During transmission the gaps at e and f spark over. The impedance of the gaps when struck is of the order of a few ohms—effectively a short circuit. The gap e prevents too much

transmitter power entering the receiver. The impedances of the gaps e and f are transformed up to effective open circuits at c and d by the $\lambda/4$ lines. Thus, owing to the $\lambda/2$ lines between a and c and between b and d, looking down at a one sees an open circuit and also looking down at b. The flow of transmitter power is thus uninterrupted. The transmission lines ch and dg across which the gaps are connected have a wide spacing and high characteristic impedance. This helps to develop the high voltage necessary to strike the gaps. It also gives a good $\lambda/4$ transformation from c and f to c and d, so that the small impedance of the gap when struck, is transformed up to as high an impedance as possible (so as to be effectively an open circuit). The limit in widening-out the spacing of the lines is set by the fact that they tend to radiate when too far apart.

72. During reception the gaps are open. Looking down at d and also at b we see a short circuit due to $\lambda/2$ transformations from the short at g. Looking to the right at b we see the transmitter and it is thus effectively short circuited at b. Looking to the right at a we see an open circuit due to the $\lambda/4$ (or $3\lambda/4$) line a b with a short at b. Looking down at a we see the receiver. The received signal from the aerial therefore passes down to the receiver and does not enter the high impedance seen looking to right at a. The gap e is often called the receiver gap, and f the transmitter gap.

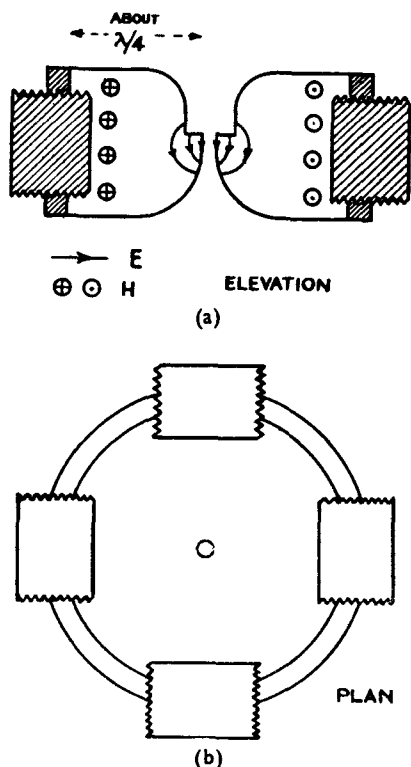


Fig. 78.—Rhumbatron

Waveguide common T and R

73. Ordinary spark gaps could not be used across waveguides since in the non-sparked-over condition they would present considerable reactance, i.e. cause reflection of the wave. The *soft rhumbatron gas switch* is used. A rhumbatron is a metallic resonant cavity as shown in fig. 78. It can be regarded as a number of $\lambda/4$ shorted transmission lines arranged star fashion. During the inductive part of the oscillatory cycle there is a strong magnetic field at the outer part of the rhumbatron. During the capacitive part of the oscillatory cycle there is strong electric field across the lips at the centre. The cavity is tuned to resonance by shortening the lengths of the "transmission lines" and this is done roughly in two or more places by screw plungers. Inductive couplings is obtained by a loop fig. 79; or by *window coupling* through a hole in the side of the rhumbatron as shown in fig. 80.

74. In order that the rhumbatron shall spark over at the lips, the centre part must be fitted into a glass envelope enclosing gas at reduced pressure. Water vapour is used since

the discharge stops quickly in this gas as soon as the transmitter goes off. In order that the discharge shall start between the lips without delay, an auxiliary discharge is run from the keep-alive electrode (fig. 81). This auxiliary discharge is outside the lips but close to them and supplies ions necessary for prompt starting of the main discharge.

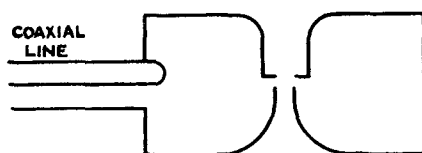


Fig. 79.—Loop coupling

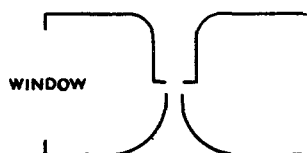


Fig. 80.—Window coupling

75. The soft rhumbatrons are connected to a rectangular waveguide through windows. Since the window is at the part of the rhumbatron where there is magnetic field, the device must be coupled to a part of the guide where there is also magnetic field, i.e. on the narrow side where the longitudinal magnetic field is strong. The arrangement is therefore as shown in fig. 82.

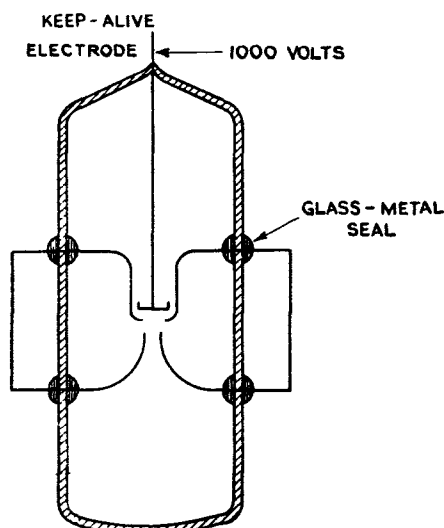


Fig. 81.—T-R cell

76. When the transmitter is on, a discharge occurs across the lips of the rhumbatrons. This destroys the resonance and there can be no large electric or magnetic field inside the rhumbatrons. There is thus no tendency for magnetic field to enter through the windows and transmission of power to the aerial is undisturbed. The receiver is coupled into one of the rhumbatrons by a loop, and no power goes to the receiver. In fact, of course, a little power enters the rhumbatrons to maintain the discharge and a very small amount of power goes to the receiver. There is also a "spike" of power into the receiver at the beginning of the transmitter pulse when the discharge in the rhumbatron is building up.

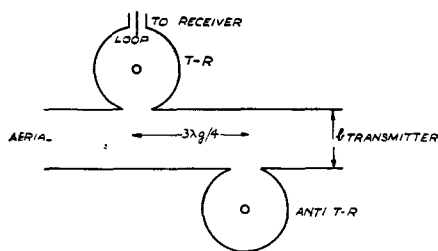


Fig. 82.—Waveguide T-R

77. During reception, the rhumbatrons resonate and accept magnetic field from the guide. Power flows from the aerial to the receiver *via* the rhumbatron and loop. The rhumbatron near the transmitter, by interrupting the magnetic field, prevents power going into the transmitter. It acts like a short circuit across the centre of the guide or an open circuit $3\lambda_g/4$ away at the receiver rhumbatron.

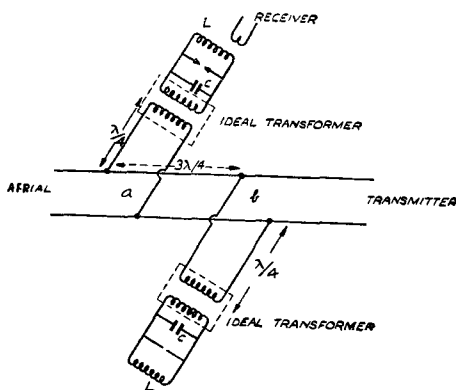


Fig. 83.—Equivalent circuit of waveguide T-R

78. The action can probably be seen more easily by replacing the guide by a parallel strip transmission line with $\lambda/4$ shorting stubs as in fig. 51. The window is represented by an ideal transformer and the rhumbatrons by resonant L-C circuits. This yields fig. 83. During transmission the gaps are struck and the L-C circuits are shorted. These shorts are transformed to shorts at the end of the $\lambda/4$ stubs and give open circuits at a and b. During reception the L-C circuit nearest the transmitter is a high impedance and after transforming gives a short at b and an open-circuit looking to the right at a. Looking up at a, the receiver is seen transformed up into the L-C circuit through the ideal transformer, and down through the $\lambda/4$ stub.

79. The rhumbatron near the transmitter is often called the *anti-T-R* switch and the receiver rhumbatron the *T-R* switch.

CHAPTER 4

CENTIMETRE VALVES

LIST OF CONTENTS

	<i>Para.</i>		<i>Para.</i>
High frequency oscillators—		Maintenance of oscillation in	
Introduction	1	magnetron	48
Transit time in triodes	2	Performance of magnetrons	54
Circuit limitations in triode oscillators	3	Magnetron pulling and frequency	
Use of special triodes at micro-wave-		splitting	56
lengths	9	Crystals—	
New types of valves	11	Introduction	66
The rhumbatron or tuned cavity	12	Description of crystal	67
Klystron oscillators	21	Classification of crystals	70
Low voltage reflector klystrons	37	Crystal mixers	71
Magnetrons	40	Standing wave detectors	76
Modes of oscillation in magnetron			
block	43		

LIST OF ILLUSTRATIONS

	<i>Fig.</i>		<i>Fig.</i>
Equivalent circuit of triode	1	Multi-resonator magnetron shown in cross-	
Ultra-audion oscillator	2	section	16
Reactance of a transmission line	3	High frequency field for π -mode in eight-	
Lecher bar tuned circuit	4	segment magnetron	17
Co-axial line oscillator	5	Strapping of eight-segment magnetron	18
Rhumbatron	6	Field conditions in anode-cathode space	19
Electro-magnetic fields in the cavity	7	Equivalent circuit of magnetron	20
S- and X-band rhumbatrons	8	Equivalent circuit of magnetron and trans-	
Cavity and electron gun	9	mission line	21
Relation between electron velocity and time	10	Magnetron frequency	22
Time/distance curves for electrons	11	Crystal	23
Reflector and double cavity klystron	12	DC characteristic of crystal	24
American low voltage klystron	13	S-band mixer	25
Klystron response curves	14	S-band waveguide mixer	26
Multi-resonator magnetron with lid			
removed	15		

CHAPTER 4

CENTIMETRE VALVES

HIGH-FREQUENCY OSCILLATORS

Introduction

1. At high frequencies normal triode oscillators have definite limitations, the most important of which are, firstly, transit time of electrons from cathode to anode, and secondly, circuit limitations.

Transit time in triodes

2. At low frequencies the anode current is in phase with the effective grid voltage, and oscillations are maintained by feed-back from anode to grid circuit. But at high frequencies, the time which the electrons take to travel from cathode to anode is comparable with the period of oscillation. This means that the grid voltage will change during the electron's transit with the result that the anode current is no longer in phase with grid voltage. Thus, the feed-back falls off and efficiency decreases.

Circuit limitations in triode oscillators

3. At low frequencies the inter-electrode capacities and lead inductances are negligible, but as frequency increases they become more and more important. The diagram, fig. 1, shows these capacities and inductances. The frequency of oscillation depends on the inductance and capacity of the tuned circuit used:—

$$f = \frac{1}{2\pi \sqrt{LC}}$$

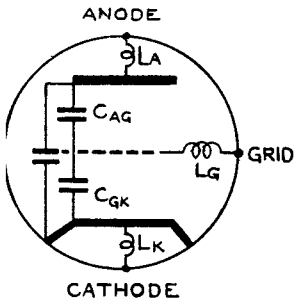


Fig. 1.—Equivalent circuit of triode

Thus, the maximum frequency will be obtained when L and C have the smallest possible values and that will be when there is no external tuned circuit, *i.e.*:—

$$\text{when } L = L_A + L_G$$

$$\text{capacity } C = C_{AG} + \frac{C_{AK} - C_{GK}}{C_{AK} + C_{GK}}$$

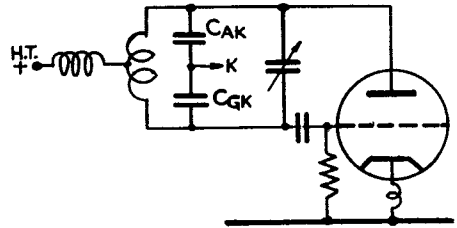


Fig. 2.—Ultra-audion oscillator

4. The ultra-audion is a type of Colpitts oscillator in which the inter-electrode capacities form part of the tuned circuit (*see* fig. 2). It will oscillate at frequencies up to 200 Mc/s.

5. The limitation of frequency by anode and grid leads (L_A and L_G) can be overcome to a great extent by making them form part of a transmission line system.

6. The input impedance of a piece of transmission line varies with its length, as shown in the graph in fig. 3. Thus, the coil and condenser of the tuned circuit can be replaced by appropriate lengths of transmission line or *Lecher bars*.

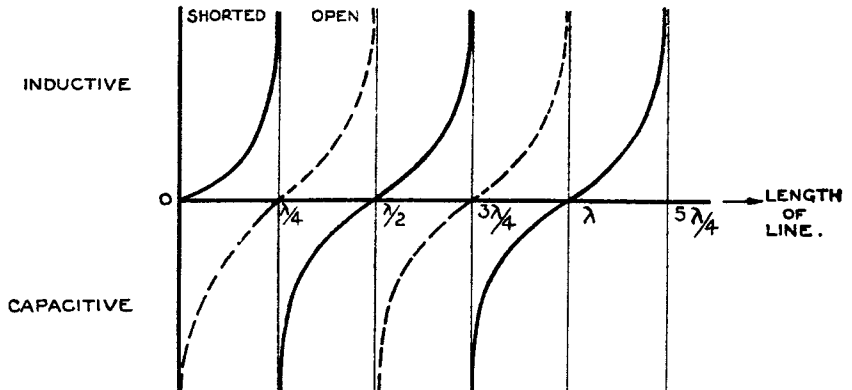


Fig. 3.—Reactance of a transmission line

7. Oscillators using Lecher bar tuning can be used up to frequencies of 600 Mc/s, and a diagram of one is shown in fig. 4. Here the length of the line includes the length of lead inside the valve itself, and the total length is less than $\lambda/4$, i.e., equivalent to an inductance. This inductance, which may be varied by altering the length of line, forms the tuned circuit along with the inter-electrode capacities of the valve.

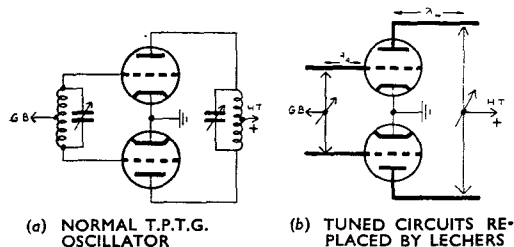


Fig. 4.—Lecher bar tuned circuit

8. To maintain oscillations in a tuned circuit energy must be supplied to make up for resistive and radiation loss. At frequencies above 600 Mc/s Lecher bar tuned oscillators become inefficient because so much energy is lost by radiation from the unscreened Lecher wires.

Use of special triodes at micro-wave-lengths

9. Special valves have been designed which use co-axial lines instead of parallel wires. Since the HF field is confined inside the outer conductor of the co-axial the radiation loss is eliminated. In this type of valve the valve itself fits into a co-axial line system as shown in fig. 5. The external circuit consists of two co-axial lines, one inside the other, formed by three concentric tubes. The outer tube, connected to the anode, and the middle one, connected to the grid, form the anode-grid co-axial, whilst the inner tube, connected to the cathode, along with the middle tube, this time

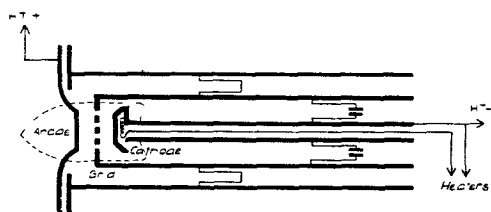


Fig. 5.—Co-axial line oscillator

acting as an outer conductor, form the cathode-grid co-axial. The frequency of the oscillation depends on the valve capacitances and the inductances due to the lengths of lines. These inductances can be altered by varying the lengths of co-axial, and thus the valve can be tuned by movable shorting pistons.

10. Valves using lecher bar tuning can be used up to frequencies of about 600 Mc/s, i.e., wavelength of 50 cms. At frequencies above that, there is too much radiation loss. Valves using a co-axial system, as described above, will oscillate at frequencies of 3,000 Mc/s, i.e., wavelengths of the order of 10 cms.

New types of valves

11. Before these special high frequency triodes were developed a completely different type of UHF oscillator was designed. This type of valve, which includes the klystron and multi-resonator magnetron, has two characteristics:—

- (i) The tuned circuit is an integral part of the valve itself.
- (ii) Use is made of the fact that an electron will travel a definite distance in one cycle.

The rhumbatron or tuned cavity

12. To increase the resonant frequency of a tuned circuit it is necessary to decrease the L and C of the circuit. If the coil is reduced to one turn of wire and a small condenser is connected across it, as in fig. 6 (a), then this tuned circuit will resonate at a very high frequency. But the losses will be very high due to the fact that the loop is unscreened and so will radiate. Also the skin effect at this high frequency will be considerable.

13. If the circuit in fig. 6 (a) is rotated about the axis 'YY', a solid figure is obtained as shown in fig. 6 (b). This solid figure is, in fact, made up of a large number (n , say) of tuned circuits in parallel. Thus the total inductance will be L/n and total capacity nC .

14. The solid figure, which is called a resonant cavity or *rhumbatron*, will therefore have the same resonant frequency as the tuned circuit.

$$f = \frac{1}{2\pi \sqrt{\frac{L}{n} \cdot nC}} = \frac{1}{2\pi \sqrt{LC}}$$

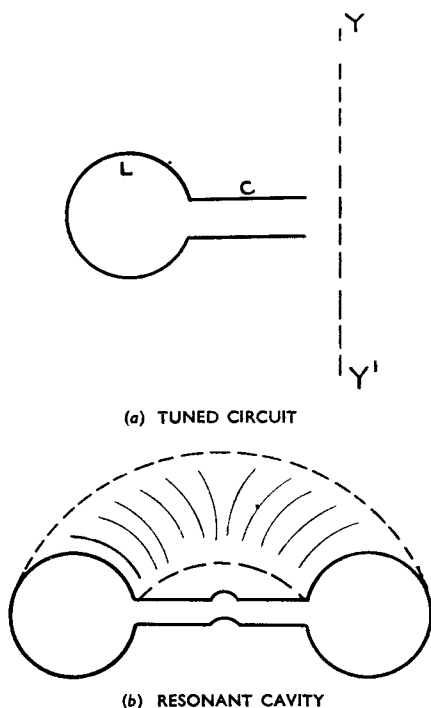


Fig. 6.—Rhumbatron

But, since the oscillations will now exist in an enclosed space, there will be no radiation. Again, since the current has a large surface area over which to flow, the skin effect, *i.e.*, resistive loss, will be reduced. Thus the *Q* of the cavity which is defined as

$$2\pi \times \frac{\text{Energy stored}}{\text{Energy dissipated}}$$

is very high—of the order of 10^4

15. In a klystron oscillator a resonant cavity of this sort takes the place of the tuned circuit and is embodied in the valve itself.

16. To examine the electro-magnetic fields which are set up in the cavity, take a section of the cavity and assume that the condenser has been charged by some means (*see* fig. 7). There will be an electric field between the plates of the condenser. The condenser is discharged by a current flowing round the loop of wire which, in turn, sets up a magnetic field. This magnetic field is at a maximum when the condenser has been completely discharged, *i.e.*, at time of zero electric field.

17. As in any tuned circuit, the inertia of the coil keeps the current flowing after the condenser has been discharged, and the condenser will be charged up again, this time in the opposite sense. Thus the electric field will be reversed, as will be the discharging current

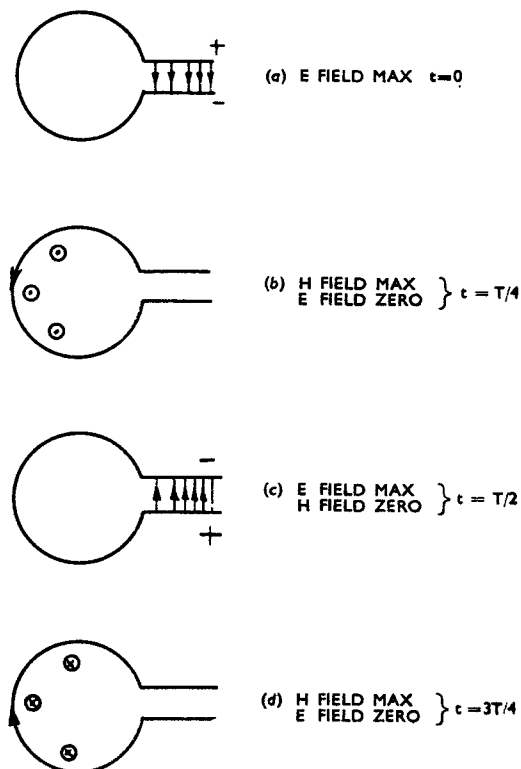


Fig. 7.—Electro-magnetic fields in the cavity

flow, which gives rise to the magnetic field. The whole cycle of oscillation can be seen in fig. 7. Since, in fact, we are not considering one turn of wire, but a solid figure, then the magnetic lines of force will not exist outside the cavity, but only inside and as closed loops round the axis of the cavity.

18. The theoretical form of cavity, as described above, has been slightly altered in practice, mainly for manufacturing purposes. Typical rhumbatrons are shown in fig. 8.

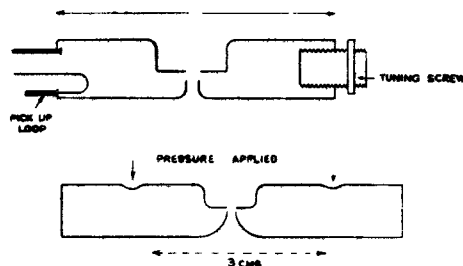


Fig. 8.—S- and X-band rhumbatrons

19. The frequency at which the cavity oscillates depends on the values of the *L* and *C* (*see* fig. 6 (a)), that is to say, on the physical size of the cavity, so that the frequency of the

oscillations can be varied by altering the size. S-band oscillators are tuned by means of three tuning screws in the side of the cavity. One of these is shown in fig. 8 (a). When these plungers are screwed into the cavity, they are decreasing the inductance and therefore increasing the natural frequency of the cavity. Conversely, screwing the plungers out decreases the frequency to a small extent. A cavity whose natural frequency is in the X-band is, of course, physically smaller and is tuned differently, since tuning screws would tend to be clumsy. Tuning is accomplished by applying pressure to the cavity by means of a ring (see fig. 8 (b)). As the top and bottom plates are pressed closer together, the capacity is thereby increased and the frequency decreased.

20. In all cavities of this type the UHF energy is taken out by means of a pick-up loop in the side of the cavity. The magnetic loops of force round the cavity thread through this loop and induce currents in it. For maximum coupling, the loop should be in a plane at right angles to the magnetic field and parallel to the electric field. Coupling can be reduced by rotating the loop from this maximum position, thus causing fewer magnetic lines of force to thread through it.

Klystron oscillators

21. Oscillations are excited in a cavity by allowing a stream of electrons in *bunches* at the natural frequency of the cavity to pass between the lips. Energy must be supplied to the fields to maintain oscillation, and this is effected by arranging that the electrons, as they pass through the lips of the cavity, give up some of their energy to the field. There are two types of klystron: (a) double cavity klystron, and (b) the reflector klystron.

Double cavity klystron

22. Consider, as in fig. 9, a resonant cavity along with an electron gun assembly. The cavity is positive with respect to the cathode, thus when electrons are emitted they are

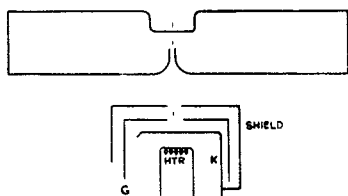


Fig. 9.—Cavity and electron gun

attracted towards the cavity and pass through the lips of the cavity.

23. Now, if oscillations exist in the cavity they will affect the electrons as they pass through. For example, if at the instant when an electron arrives at the cavity, the field is as shown in fig. 7 (a), then the electron will be travelling towards a positively charged plate, and so its velocity will be increased. Similarly, the velocity of an electron arriving half a period later will be decreased. It must be remembered that the charge is only inside the cavity, *i.e.*, on the inner side of the plates. An electron arriving at a time half-way between these two would travel on with its velocity unchanged, since there is no charge on the plates at this time (see fig. 7 (b)).

24. Suppose that the cathode-cavity is such that all the electrons have a velocity ' v ' when they reach the cavity, and suppose that oscillations have been shock-excited in the cavity. It has been seen that an electron arriving at the cavity at time $t=0$ (see fig. 10), will have its velocity increased—by Δv say. An electron arriving at $t=T/4$ passes through with its velocity unchanged, *i.e.*, v , and one arriving at $t=T/2$ has its velocity decreased—by Δv .

25. When the stream of electrons has passed through the cavity, the velocities of the electrons will have been altered. This is called *velocity modulation*, and for efficient velocity modulation the lips of the cavity must be close together so that the field does not change during transit time of electrons.

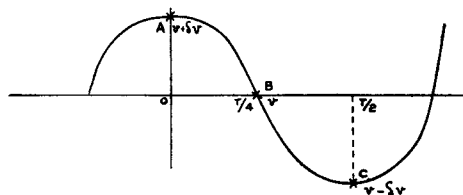


Fig. 10.—Relation between electron velocity and time

26. Fig. 11 is a graph plotting the distance travelled by the electrons after they have passed through the cavity, against time. Thus the slopes of the lines represent the velocities of the different electrons. It can be seen that at some distance d from the origin, *i.e.*, from the cavity, the lines cross. This means that the electrons arrive in bunches at this point, at the resonant frequency of the cavity, the faster-moving electrons having caught up with the slower-moving ones. This place where

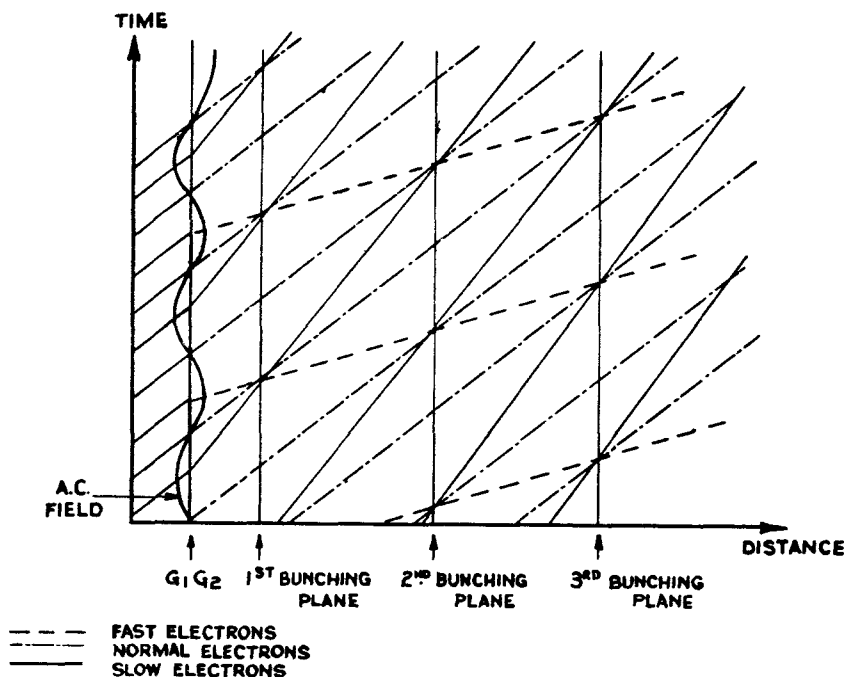


Fig. 11.—Time/distance curves for electrons

the electrons bunch together is called a *bunching plane*, and there are a number of these planes, three being shown in the diagram.

27. Now, if a second cavity, whose resonant frequency is the same as the first, is placed at a bunching plane of the first cavity, then the electrons will pass through the lips in bunches at its natural frequency and thus excite oscillations in it.

28. The first cavity, which causes the electron stream to be velocity modulated is called the *buncher*, and the second the *catcher*. There is feed-back of energy from catcher to buncher by means of pick-up loops or by coupling through a hole, and thus fairly large oscillations are built up. An electrode above the cavity and slightly positive with respect to the cavity collects the electrons after they have passed through.

29. This type of klystron is called a double-cavity klystron, and a typical example is shown diagrammatically in fig. 12 (a). The pick-up loop and tuning screws are as described above. Both cavities must be tuned to the same frequency. It has an output of approximately 10 watts CW or up to 100 watts if water cooled, but its efficiency is very low, mainly due to the fact that the catcher cannot be placed close enough to the buncher for mechanical reasons. It is not possible then to use the first bunching

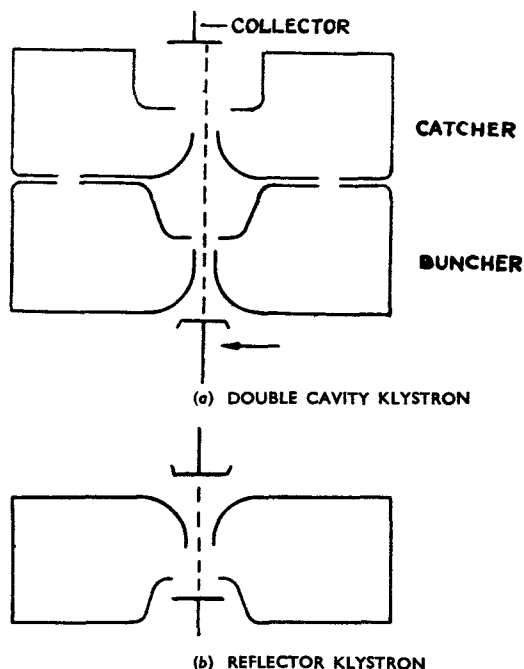


Fig. 12.—Reflector and double cavity klystron

plane, but some higher order plane must be chosen. Now, as the electrons move away from the buncher, they tend to interact and

repel each other, thus demodulating the stream, the result being that they no longer bunch together so efficiently as they get further and further away from the buncher.

Reflector klystron

30. The reflector klystron is the more common type, and is mainly used as a local oscillator or as a test set source since its power output is of the order of 50–200 milliwatts CW.

31. This klystron, illustrated in fig. 12 (b), has only one cavity and an electrode called the reflector which is at a highly negative potential with respect to the cavity. Thus there is a very strong field between cavity and reflector. Again the electron stream is velocity-modulated as it passes through the cavity. After passing through, electrons meet the retarding field due to the negative potential on the reflector, and are gradually slowed down. At some point the electron velocity becomes zero, after which the electrons acquire a velocity in the opposite direction and return to the cavity.

32. A convenient analogy may be drawn with stones thrown into the air, the force of gravity corresponding to the retarding force due to the reflector. Stones thrown upwards with greatest initial velocity will go furthest up, and so will stay longest in the air. A stone thrown up later with a smaller velocity will not remain so long in the air. So that by arranging their initial velocities and difference in time between their leaving the ground, they can be made to arrive back at earth at the same time.

33. Consider the three electrons indicated in fig. 9, *i.e.*, those whose velocities on leaving the cavity are $v + \Delta v$, v and $v - \Delta v$. The first will go furthest towards the reflector before being repelled back, and so will be in the retarding space the longest. The third electron, which leaves half a period later, will be in the retarding space for the shortest time. By adjusting the reflector voltage, it can be arranged that these three electrons arrive back in a bunch at the cavity.

34. Thus, in a reflector klystron, the electron stream is velocity-modulated as it passes through the cavity; we must assume that oscillations have been shock-excited in the first place. As they come back, having been reflected, they are in bunches at the frequency of the cavity and so will excite oscillations.

35. But to maintain oscillations energy must be supplied to the field from the electrons. Now, if an electron is travelling in the same direction as the field in which it is in, then it will give up some of its energy to the field. Whereas, if it is travelling against the field, it will take energy from it. This can be seen again in fig. 10. At A the field is downwards, *i.e.*, from cavity to cathode, and the electron is moving upwards against the field. The electron's velocity has increased, therefore it must have gained momentum at the expense of the field. At C, the electron travelling in the same direction as the field, has lost momentum and so must have given energy to the field.

36. Thus, it should be arranged that, during the time in which electron C is in the retarding space, the field in the cavity reverses. So that when this electron, in a bunch with other electrons, returns to the cavity it is again travelling in the same direction as the field, *i.e.*, downwards, and the electrons give up their energy to the field. This means that the time which electron C is in the retarding space must be an odd number of half-periods, *i.e.*, $T/2$, $3T/2$, $5T/2$, etc. It can be arranged by correct voltages on the various electrodes. In the normal type of British manufactured klystron, the cavity is earthed, the cathode at $-1,600V$ approximately, and the reflector at $-2kV$. The reflector electrode is placed very close to the cavity to prevent any demodulation of the electron stream. The tuning controls and pick-up loop are the same as in the double cavity klystron.

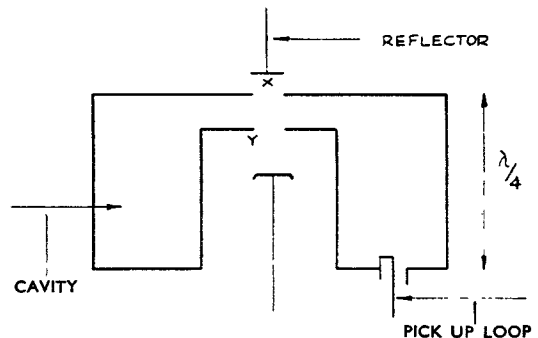


Fig. 13.—American low voltage klystron

Low voltage reflector klystrons

37. American low-voltage reflector klystrons are slightly different from those described above, but they work on exactly the same principle. The cavity is usually regarded as a resonant length of co-axial transmission line,

i.e., a quarter wavelength long shorted at one end (see fig. 13). The resonant frequency will therefore depend fundamentally on the length of the line. The cavity is tuned in rather the same way as the X-band cavity described above. The plates X and Y can be squeezed closer together, thus adding more capacity across the line and so effectively lengthening it and decreasing the resonant frequency. The output loop is continued as a probe through the valve base so that, with the valve sitting on top of the guide, an H01 wave can be propagated directly into the guide.

38. The main characteristic of these low voltage klystrons is that they have a low Q cavity and consequently the response curve of the valve is fairly flat. Thus, the conditions for the valve to oscillate, principally the reflector voltage, are not highly critical. In fact, the valve will oscillate over a definite range of reflector voltage and frequency, the frequency increasing with increase of negative reflector voltage. The diagram (fig. 14) shows response curves of (a) high Q cavity oscillators and (b) low Q cavity oscillators.

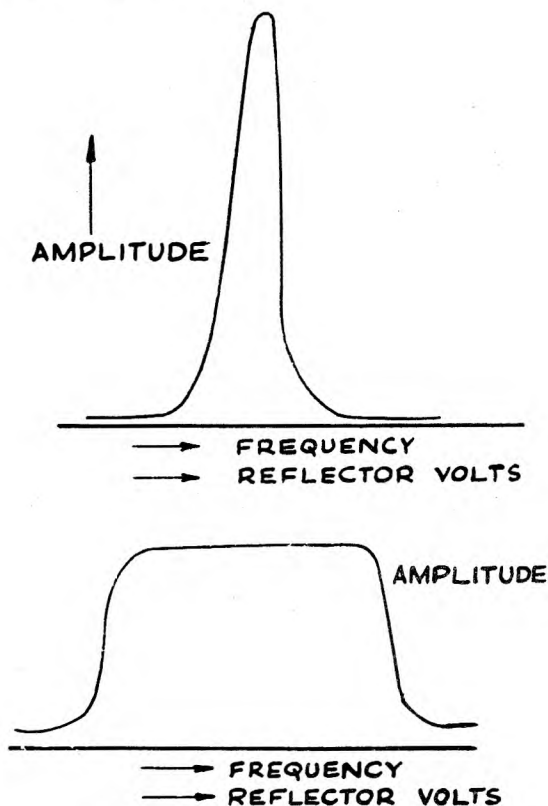


Fig. 14.—Klystron response curves

39. Thus these low voltage klystrons (cathode 0v, cavity +300v, reflector -90v) can be made to oscillate over a band of frequencies without altering the cavity dimensions. They are therefore used as local oscillators in equipments which require automatic frequency control.

Magnetrons

40. The magnetron is, so far, the only valve capable of giving high peak power output at ultra high frequencies. The type described here is the multi-resonator magnetron and not the split-anode type which will not oscillate efficiently at frequencies above 600 Mc/s.

41. In figs. 15 and 16 there are shown two cross-sections of the multi-resonator magnetron. The anode is a metal block round a cylindrical, oxide-coated metal cathode. As in the klystron, the oscillatory circuit is embodied in the valve itself. In this case, there are a number of such circuits, called *resonators*, cut out of the anode block. Each resonator consists of a slot and a hole in the anode, the slot acting as a condenser and the hole as an inductance. If, in some way, this condenser is charged up and the LC circuit left to itself,

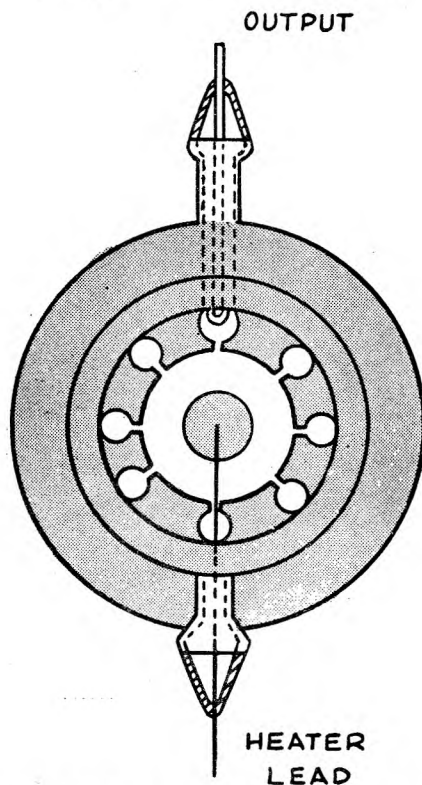


Fig. 15.—Multi-resonator magnetron with lid removed

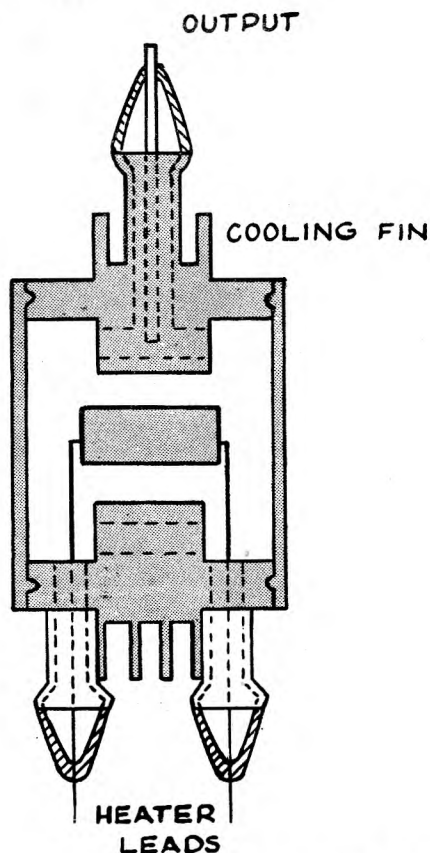


Fig. 16.—Multi-resonator magnetron shown in cross-section

it will oscillate at its natural frequency, and the RF oscillations can be picked up by a loop placed in the hole.

42. However, each oscillatory circuit cannot be considered separately since they are all coupled together. There is mutual inductance between the holes, since magnetic lines of force emerging from the top of one hole must pass down through adjacent holes in order to form closed loops. There is also capacitive coupling due to the capacities between the anode and end plates. Due to this coupling, therefore, the resonators do not oscillate independently but together. They can oscillate together in a number of different ways or modes, of which only the simplest will be considered.

Modes of oscillation in magnetron block

43. It is more convenient to draw a magnetron "rolled out", that is, the cathode and anode as

straight lines. The diagram (fig. 17) shows the potentials on the anode segments when the magnetron is oscillating in the simplest mode. Alternate segments are at equal and opposite potentials so that adjacent resonators are oscillating with 180 deg. phase difference. This mode is therefore called the π -mode, and is the one in which nearly all magnetrons work.

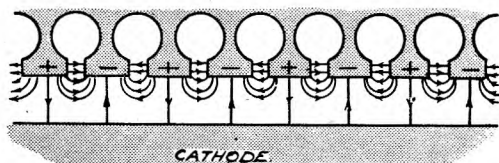


Fig. 17.—High frequency field for π -mode in eight-segment magnetron

44. The electric and magnetic fields in the resonators are the same as shown in fig. 7 (klystron), but in the case of the magnetron, the electric fields extend into the anode-cathode space where the electrons will be moving (see fig. 17). This *end effect* is important in considering how a magnetron works.

45. The frequencies of different modes of oscillation are practically the same, with the result that the magnetron will tend to jump from one mode to another. To prevent this a method of *strapping* (see fig. 18) has been used. This consists of a number of wires connecting alternate anode segments. The effect of these wires or straps is to separate out the frequencies of the modes so that for given conditions the magnetron will oscillate in one mode and one mode only.

46. Straps are also used for tuning magnetrons since capacity exists between them and the anode segment over which they are passing. This capacity is shared between the two resonators on either side of the segment, thereby decreasing their natural frequency. Thus by moving the straps up and down, *i.e.*, decreasing or increasing the capacitive effect, the frequency of the oscillators can be varied.

47. The fundamental frequency at which a magnetron oscillates depends mainly, of course, on the physical size of the slots and holes, but also to some extent on the amount and nature of the coupling between the resonators.

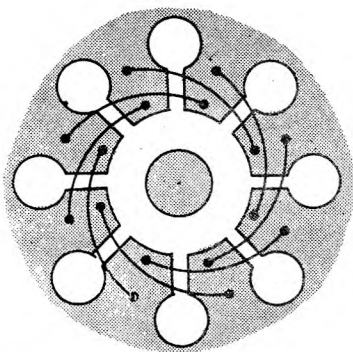


Fig. 18.—Strapping of eight-segment magnetron

Maintenance of oscillations in magnetron

48. The operating conditions for a magnetron are (a) a potential of the order of 10kV between anode and cathode; (b) a magnetic field of about 1,500 gauss parallel to the axis of the cathode. This is obtained by placing the magnetron between the poles of a permanent magnet.

49. When an electron leaves the cathode it is attracted across to the anode, but the presence of the magnetic field gives it a sideways motion, so that it moves round the cathode and out to the anode at the same time. Electrons moving towards the anode segments induce charges on them and thus charge up the condensers of the oscillatory circuits. Each of these circuits will therefore swing into oscillation, but energy will have to be continually supplied to them if they are to keep oscillating.

50. Now, during its transit from cathode to anode, an electron will pass a number of slots, and, as described above, some of the RF field has spread out from these slots into the anode-cathode space. Thus the electron will have passed through a number of RF fields before it reaches the anode. If it can be arranged

that all the electrons which reach the anode are travelling in the same direction as the RF field in which they find themselves, then they will give up some of their energy to the fields and thus maintain them.

51. In fig. 19, due to the magnetic field B and electric field E , there is a force on the electrons to the right. Suppose the first RF field in which an electron finds itself is in the same direction as that in which it is moving, then the electron gives energy to this field. Now, if the magnetron is oscillating in a π -mode, then the RF fields emerging from adjacent slots are 180 deg. out of phase. Thus, if the electron is to give up some energy to the adjacent RF field, then this field must have reversed when the electron reaches it. This means that the time which an electron takes to travel from one slot to the next must be half a period.

52. The speed velocity of the electrons round the magnetron depends on the values of the E and B fields, and these values are linked up with the period of oscillation of the magnetron, as has just been shown. Therefore, if a magnetron is to oscillate at a given frequency, the correct values of electric and magnetic fields must be provided.

53. So far, the electrons which would take energy from the RF fields have been ignored. In practice these electrons never reach the anode, but return to the cathode almost as soon as they are emitted. In some magnetrons, these electrons hitting the cathode are sufficient to keep it heated and, once the magnetron is oscillating, the heater is switched off.

Performance of magnetrons

54. The RF fields in the slot and hole circuits build up to any high values, and the peak output power of a CV64 S-band magnetron is 40kW approximately. The energy is taken from the magnetron by a loop in one of the resonators. The loop joins a co-axial line which is brought out of the magnetron through a metal/glass seal.

55. Magnetrons are capable, therefore, of producing very short pulses of high peak power, but in their present form are incapable of operating on a continuous basis. This is because firstly, a very large number of electrons is required to produce conditions for oscillation and oxide coated cathodes are incapable of supplying sufficient continuously, and secondly, since magnetrons are 30–40 per cent. efficient, it is necessary to dissipate large quantities of wasted power. A large surface

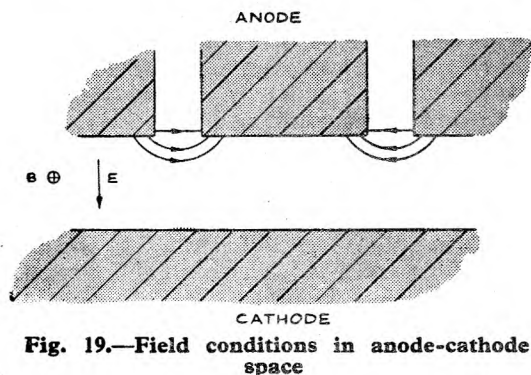


Fig. 19.—Field conditions in anode-cathode space

area, usually in the form of cooling fins, is supplied, such that this wasted power in the form of heat may be readily dissipated by air blast. It is important that the temperature of the magnetron anode be kept constant, since thermal expansion will produce changes in the physical sizes of the resonators, and subsequent changes in frequency.

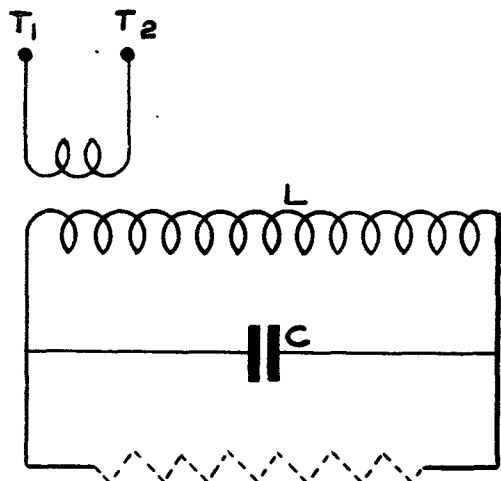


Fig. 20.—Equivalent circuit of magnetron

Magnetron pulling and frequency splitting

56. A multi-resonator magnetron which has eight resonators, each consisting of an inductance L and capacitance C , is equivalent to a single tuned circuit whose inductance is roughly $L/8$ and capacitance $8C$. Thus the magnetron with its output loop can be represented as in fig. 20 (the negative resistance across the tuned circuit representing the electrons which supply energy to maintain oscillations).

57. Any reactance connected across the terminals T_1 and T_2 will be reflected back into this tuned circuit and therefore alter its resonant frequency. The magnetron transmitter has the aerial directly coupled to its *tank circuit*, the result being that its operating frequency depends on the load into which the valve is working.

58. Since most magnetrons are untuneable, this property may be utilised by placing a reactive stub near the valve output. As the length of this stub is altered, different resistances will be reflected back into the tuned circuit of the magnetron thus varying its operating frequency.

59. The question of pulling becomes more serious when the magnetron is connected to the aerial by a long length of cable or waveguide. If the aerial terminates the cable or guide in its characteristic impedance, then no standing waves exist and the magnetron will work normally. But this termination is not easy to obtain in practice, and, if the aerial is not matched to the cable, standing waves will be set up and a certain impedance will exist across the terminals X and Y in fig. 21. This impedance will have a reactive term which depends on the length ' l ' of the feeder.

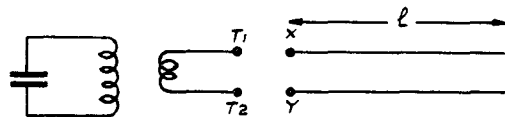


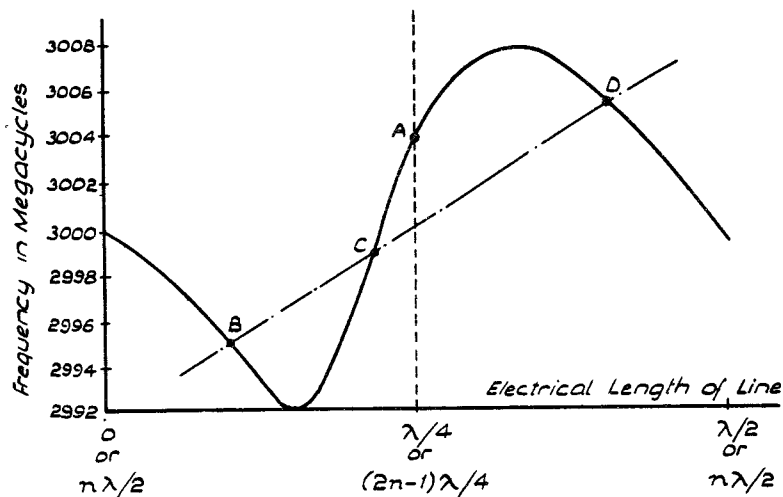
Fig. 21.—Equivalent circuit of magnetron and transmission line

60. If l is less than quarter of a wavelength, then the reactive term will be capacitive. This capacitance presented at T_1T_2 , after being reflected through the transformer, will appear as additional inductance in the L - C circuit, thus decreasing its natural frequency. If l is between $\lambda/4$ and $\lambda/2$, an inductance as presented at T_1T_2 which, on reflection, appears as additional capacity in the tuned circuit thus increasing its resonant frequency. The graph of operating frequency against electrical length is therefore as shown in fig. 22 (full line).

61. So far, only the electrical length of the line has been considered, and now the variation of this electrical length with frequency must be found. If the line is physically short—say 2.5 cms., then at a frequency of 3,000 Mc/s its electrical length is $\lambda/4$ and very nearly so at neighbouring frequencies. But if the line is long (say 1002.5 cms.) at 3,000 Mc/s, its electrical length is $100\frac{1}{4}\lambda$, at 2,995 Mc/s it is $(100\frac{1}{4} - \frac{1}{8})\lambda$, and at 3,005 Mc/s it is $(100\frac{1}{4} + \frac{1}{8})\lambda$.

62. Thus, if the length of line is short, the graph of electrical length of line against frequency will be a vertical line (dotted line in fig. 22), but if it is a long feeder, the graph will be the dot-dash line, *i.e.*, electrical length increasing with frequency.

63. Since the frequency of operation is a function of electrical length of line and the electrical length is, in turn, a function of



FULL CURVE—MAGNETRON FREQUENCY AS A FUNCTION OF ELECTRICAL LENGTH OF LINE, FOR 2 : 1 STANDING WAVE ON LINE

DOTTED LINE—ELECTRICAL LENGTH VERSUS FREQUENCY FOR A LINE 2.5 CM. LONG

DOT-DASH LINE—ELECTRICAL LENGTH VERSUS FREQUENCY FOR A LINE 1002.5 CM. LONG

Fig. 22.—Magnetron frequency

frequency, the magnetron will operate at the point where the full curve is intersected by the dotted line.

64. It can be seen that with a feeder length of 2.5 cms. there is only one operating point A, but with a long feeder there are three operating points, B, C, D. The magnetron will be

unstable and will tend to jump between these three frequencies.

65. This effect of multiple frequencies is called frequency splitting, and can be eliminated by keeping the length of line as short as possible and also by having no standing wave on the line.

CRYSTALS

Introduction

66. Since, at wavelengths of 10 cms. and less, amplification is impossible at the signal frequency, the first stage of the receiver is the mixer. Most centimetric equipments use a crystal as a mixer, and also as a low-level detector. Crystals are also extensively used as indicators for standing wave measurements.

Description of crystal

67. The type of crystal used is shown in fig. 23. The rectifying contact is made by placing a pointed tungsten whisker on a smooth silicon surface, when the electrons will flow from the metal to the crystal. The tungsten wire is held by the top cap and the crystal is mounted on the bottom pin. Loss-free ceramic material connects the top cap and pin.

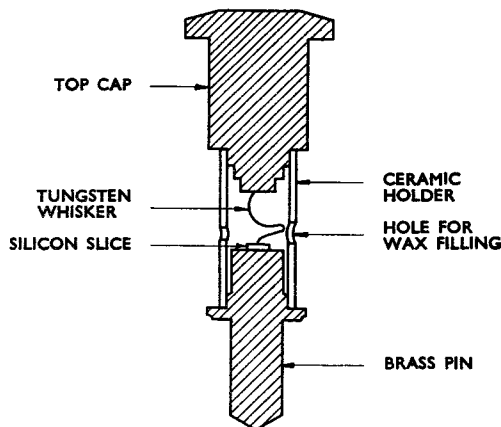


Fig. 23.—Crystal

68. The DC characteristic of a crystal is shown in fig. 24. Thus if a battery is connected across the crystal with top cap to negative then current will flow through the crystal depending on the applied volts. The resistance of the crystal will be of the order of 2-300 ohms.

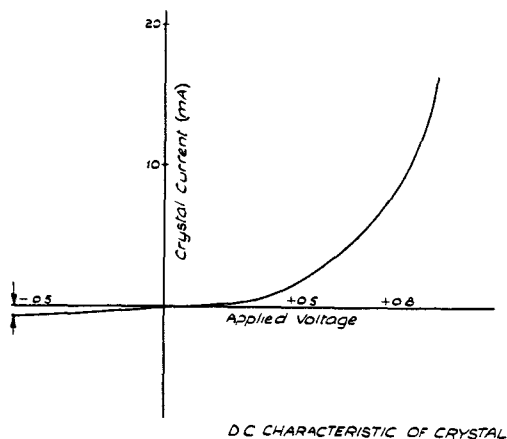


Fig. 24.—DC characteristic of crystal

69. But if the crystal is reversed only a very small current will flow as the back resistance is very high (over 3,000 ohms). The ratio of the back-to-front resistance will give an indication of crystal stability, although there is not complete correlation between this ratio and the performance of the crystal. The back-to-front ratio should be greater than 10:1.

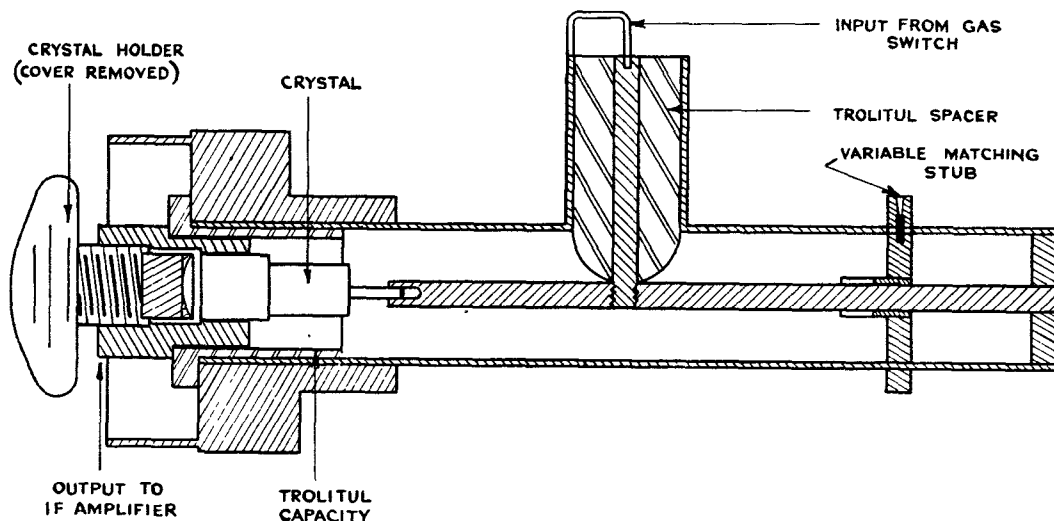
Classification of crystals

70. Crystals are classified by (a) the waveband over which they have been selected to operate, and (b) their resistance to RF burn-out.

- (i) S-band crystals have a **YELLOW** spot, and X-band crystals a **GREEN** spot painted on them.
- (ii) An efficient TR box should prevent any of the transmitted pulse getting through to the crystal. However, in practice a 'spike' of energy from the transmitter pulse does break through the TR box before the gas has had time to ionise completely. The duration of the 'spike' is about 10^{-9} seconds and the burn-out will depend on the total energy in the spike. The crystals are classified, therefore, by the spike energy required to deteriorate the crystal by a certain amount (2db). **RED** spot crystals have a high resistance to burn-out. **ORANGE** spot crystals gave a medium resistance to burn-out. Only one spot, the waveband spot, is on crystals with a low resistance to burn-out, the others having two coloured spots.

Crystal mixers

71. It has been found that the resistive component of the crystal RF impedance does not vary much among crystals, although the reactive component does. Thus it has been



possible to design a crystal holder with only one variable, to tune out the crystal reactance.

72. Fig. 25 shows an S-band crystal mixer. The crystal fits into a co-axial line whose length can be altered, thus tuning out the crystal reactance. The input from the TR cell or gas switch is fed on to the inner conductor by a stub on one side, and the local oscillator frequency, by another stub.

73. When these two frequencies, f_1 and f_2 say, are applied to the crystal which has a non-linear characteristic, a number of frequencies are produced— f_1^2 , f_2^2 , $f_1 + f_2$, $f_1 - f_2$, etc. A small trolitul capacity round the crystal presents a dead short to the high frequencies, but an appreciable reactance to the beat frequency $f_1 - f_2$. Therefore this IF can be picked off across the trolitul condenser.

74. The X-band crystal mixer has the crystal positioned across the guide so that the resistive component of its impedance is matched to the Z_0 of the line, and the reactive component is tuned out with a backing piston (see fig. 26). The crystal should be placed so that the electric field in the guide is parallel to the length of the crystal.

75. The local oscillator frequency is again

injected from the side and the IF output is taken across a mica capacity.

Standing wave detectors

76. On both S- and X-band standing wave detectors the crystal is fitted into a co-axial line very similar to the S-band mixer. A probe, which projects into the guide, is connected to the inner conductor of the co-axial. This probe picks up the RF voltage in the guide, the crystal detects it, and the DC output is taken off across a small capacity. Again, the length of the co-axial can be altered to tune out the crystal reactance.

77. The diagram of fig. 24 shows that a crystal has a square law characteristic up to a crystal current of about $20\mu\text{A}$. Thus, in the detector described here, the reading on the meter connected across the mica capacity, is the square of the applied voltage. If standing wave ratios are to be measured, the square root of the ratio of maximum to minimum meter readings should be taken.

78. Care should be taken when handling crystals as they will stand only a limited amount of mechanical shock. They can also be damaged by the passage of discharge currents when being inserted into unearthed apparatus.

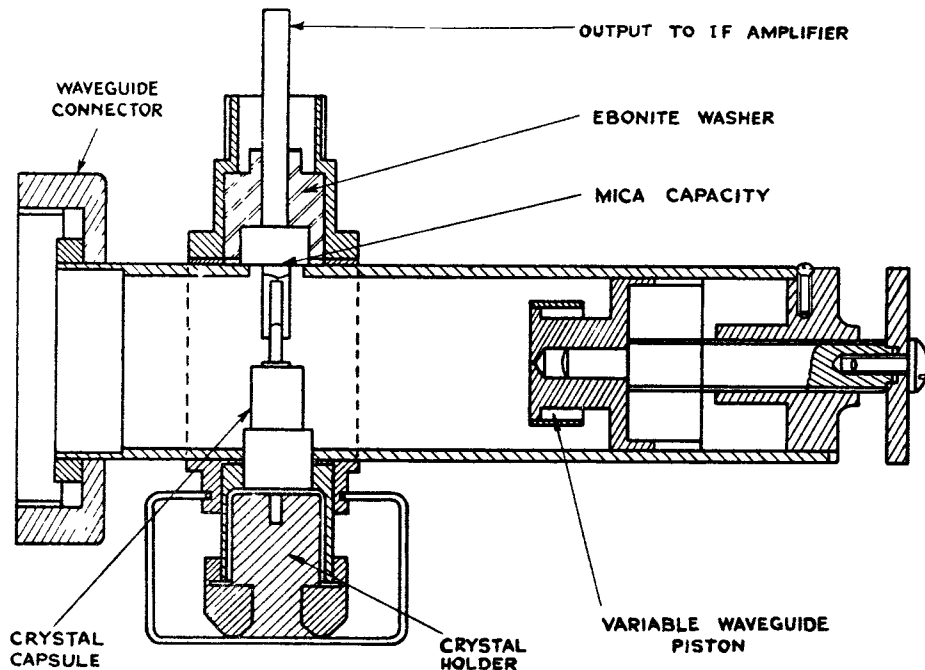


Fig. 26.—X-band waveguide mixer